Spring 2018

Spoonerisms: An Analysis of Language Processing in Light of Neurobiology

Naomi Sellers
Regis University

Spoonerisms: An Analysis of Language Processing in Light of Neurobiology

A thesis submitted to
Regis College
The Honors Program
In partial fulfillment of the requirements
For Graduation with Honors

Naomi Sellers

May 2018

Thesis written by
Naomi Sellers

Approved by

____
Thesis Advisor

____
Thesis Reader

Accepted by

____
Director, University Honors Program

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my advisor Dr. Nicholas Myklebust for such incredible guidance, for offering so much of his time and knowledge and teaching me so much along the way, for investing his interest in this quirky topic, and for his overall support throughout the process. This thesis would not exist as it is without the copious amounts of meetings and draft reviews. I have learned so much from him over these last four years in class, in thesis meetings, in simple conversation. I am truly blessed. Thank you to Dr. Amanda Miller for not only agreeing to be my thesis reader but helping me trek through my primary experiment. It would have been a major struggle to figure out all the IRB requirements alone without her guidance, and her support and encouragement helped me throughout the experiment process.

Thank you to Dr. Lara Narcisi, Dr. Catherine Kleier, and Dr. Thomas Howe for providing support and direction throughout the Honors program and through the thesis process. I would like to thank the Regis CC&IS and the Data Sciences Department for access to and use of the OpenBCI EEG equipment. Thank you to the Honors cohort and to my lovely Regis family for their encouragement, comradery, and help in all the various ways that made this thesis process possible, that made classes thought-provoking, and that filled pastimes with joy. I would like to thank my wonderful family who have supported me throughout this journey, and a special thanks to my sister for helping me develop my writing over the years and who took the time to work through my thesis with me. Fæder, iċ þē geþancie for þāra me gyfst in mīn fēore. Geþancie þē for þīn hēahlufan on frēode ond ġefēan.

Chapter 1: Introduction

What are Spoonerisms?

Language is a wonderful and curious thing. It is a process we use to communicate with others, with the world around us, and oftentimes with ourselves. We use language to interact not only with one another but as the medium for interactions across all subjects: history is the documentation or oral relation of significant interactions; literature is written interactions with different purposes that can be studied through different lenses; visual and performing arts are the physical manifestation of ideas or feelings being communicated; mathematics and scientific notations are the words and symbols we use to communicate and interact with the world around us and to describe the world interacting with itself. Language is not just a process or a tool we can use as it is interwoven into our lives in various and interesting ways. Language is a complex structure that tightly ties our humanness—culture, social structures, history, art, philosophy—to our biology, chemistry, and physicality. It can shape our identity as much as we can use it to express our identity. Language colors our perception of the world around us and our own idiolects contribute to our perspectives and ideas. Language affects the brain and our sense of self just as much as the brain and our preferences affect the language we use.

It is this curiosity of how language and the brain interact that draws me the most to study linguistics, though the two fields of study—of language and of the brain—are often separated from each other. There is often a sort of categorization applied to the study of linguistics as a whole when scholars work in isolation from other subjects, such as the separation of linguistics from neuroscience. Yet this separation is often applied just as much to the more detailed components of linguistic studies when researchers and scholars portray linguistics and language processing as categorical, as something being either true or false.

One particular example of this linguistic categorization is with the perception that language processing can be either "correct" or "errored" in a way that suggests some sort of pathology. However, a number of classes, personal studies, and personal experience have prompted me to question the perception of clear categorization of cognitive functions and behavior as binaries, like the normative versus pathological, and to explore the possibility of a continuum-based model of functions and behaviors instead. My neuroscience classes have highlighted the lack of neuroanatomical and empirical support many linguistic theories have and the disparity in cross-disciplinary communication between the two fields in terms of structure and terminology. Language processing should not be so strongly separated from neuroscience, but rather the two fields should interact and collaborate on a model that is linguistically and neuroanatomically acceptable and supported.

This is where the study of spoonerisms comes in. Have you ever experienced a time when you were speaking and you come across a word or a few that, for some reason, come out all jumbled up and mixed together? Have you ever tried to say "top hats on bald cats" but instead say "tot haps on cald bats"? This category of speech errors involving jumbled-up words occurs when one sound unit in one word switches with a sound unit in the other word in a process known as "metathesis." For example, *rats and mice* might become *mats and rice,* or *blue hats* becomes *hue blats*. This type of speech metathesis is called a “spoonerism.” While many researchers describe spoonerisms—along with other types of speech errors or productions like tongue twisters—as a complete and discrete switching of sound categories, the findings of many more recent studies suggest that spoonerisms are actually a speech error produced along a sound grading continuum anywhere between the intended and unintended sounds (Goldrick et al, 2016).

So why do spoonerisms happen? Moreover, why do mix ups like this become difficult to correct, even when you are aware of the problem? For me, my interest in this verbal slip phenomenon came when I was trying to say “Father Woody” to a friend but all I could say was *wadder foodie* no matter how hard I tried. I spoonerize quite frequently, but this bizarre instance stayed with me and I began wondering not only why do spoonerisms occur, but is there a correlation between spoonerisms and the neural structures involved in language processing? To start, I began by looking into the historical documentation of spoonerisms.

The origins of “spoonerisms” as a term to describe such speech error began in the early twentieth century with Reverend William A. Spooner, the dean and warden of New College, Oxford (Fromkin, 1973). Rev. Spooner was attributed with making verbal slips where two sounds of an intended utterance were switched, such as saying “the queer old dean” instead of “the dear old queen,” or “work is the curse of the drinking class” in place of “drink is the curse of the working class.” Though the nomenclature comes from Rev. Spooner, spoonerisms have a much more expansive history. From the intentional spoonerisms in literature by writers like Shakespeare and Shel Silverstein, to the unintentional slips by newscasters, and to the psychology studies like those of Sigmund

Freud, spoonerisms have marked their impact on the way we use our language beyond just a speech error.

Over time, spoonerisms began to attract the attention of some linguists and psychologists who started studying the phenomenon in hopes of better understanding language processing behavior. Victoria Fromkin (1973) describes the early history of spoonerisms, beginning with neurophysiologist Karl Spencer Lashley who regarded speech errors as evidence that a hierarchy of organization could account for speech behavior because the disruption of the hierarchical units would result in said speech errors. From Lashley's description, Fromkin establishes this hypothesized hierarchy's framework by explaining how a speaker is able to form a potentially infinite set of sentences by building along the hierarchy of units, from phonemes to morphemes to words to sentences. In this hierarchy, the phonetic units may have real features independent of mental grammar because voicing switches of phonetic units occur separately from the grammar. Fromkin states that the existence of speech errors also suggests we may learn morphemes as separate items from the rules for their combination, allowing us to create new words and correctly produce morphemic combinations we have never heard before. This could mean that spoonerisms may occur due to a production error in the ordering and combining morphemic processes, or that it may be an input error in the encoding process for proper morphemic combinations.

Lashley and Fromkin's theory that spoonerisms are evidence for organizational hierarchy is just one of many theories on language processing that drives spoonerism studies. There are two major categories of language production models that many of the theories fall into: modular models and connectionist models. However, despite these various studies and theories, there still is no consensus of how exactly spoonerisms are produced, why they are produced, and how they fit into a language model with explanatory adequacy.

Moreover, the bulk of the current research and literature is based on linguistic and psycholinguistic theories and methodology, so the neurolinguistic point of view is particularly deficient. Language is a biological process because biological capabilities limit the types of language processing that can occur in a real human being, and as such the study of language processing (and by extension, spoonerisms) is not relevant to only the field of linguistics or only the field of neuroscience but to both. Since the processes that linguistics and neuroscience each study are really interdependent, the language models neuroscientists use need to agree with established linguistic structures, and linguistic theories need to have a biological adequacy that is consistent with neuroscientific evidence.

Because it is a very under-researched phenomenon, spoonerism production may seem like a strange and niche study with limited application, but it is because of this deficiency in the research that spoonerisms should be studied. As earlier researchers have asserted, spoonerisms can offer nuanced insight into language processing by providing opportunity to test the neuroanatomical basis for spoonerism production (and by proxy, other verbal slips). The elicitation of spoonerisms may also reveal where along the language acquisition-to-production process this verbal slip occurs, or at least typically occurs.

The overarching desire for this research is to better understand language deficits with neurological causes as well as understanding normal language processes and the mechanisms that facilitate it. Better understanding of normal language processing could further help us understand language pathologies like dyspraxia and aphasias; spoonerisms, as well as other "benign" speech errors, may actually be a type of bridge between normative and pathological language processing as they push against the decisive separation and call into question whether these two side of a diagnosis are really binary, suggesting that perhaps language processing function exists on a continuum rather than being clearly and definitively either normative or pathological. While the purpose of this particular research is to contribute to a language model with grammatical competence that reconciles explanatory adequacy with empirical neuroanatomy, my ultimate hope is that this research may contribute information that can be used to help people who face neurological linguistic challenges.

Chapter 2: Problems and Requirements for Language Processing Theories

Language processing is a complicated activity because there are a large number of mechanisms and components involved, many of which are interdependent and integrative. For just one sentence to be produced verbally, a nonlinguistic message must somehow be converted into a signal that activates the right phonological, morphological, lexical, syntactic, and semantic information in the right sequencing order. These activated information units must then be converted into the right motor commands (while retaining the proper order) before all the necessary articulators then coordinate the specific fine motor movements with accuracy to cause particular turbulences in the airflow, all the while non-necessary articulators avoid changing the airflow's turbulence or interfering with the active articulators' movements. Then, in order for that one sentence to be understood by a listener, the sound waves produced by the turbulent air stream must enter and process through the listener's auditory system, at the end of which the speech is reconverted into linguistic information. Finally, the received linguistic information must be broken down and decoded into meaning.

The three main questions a language model must consider when attempting to map out the language-processing system are (1) how is information encoded, (2) how is language stored, and (3) how is language accessed? Understanding how language information is encoded is important because encoding is involved in language acquisition as well as in language production (including speech, sign, and written). In addition, understanding how information is encoded may give insight into whether or not any information is innate, such as a universal grammar, or if all linguistic information must be first learned. The way language is stored is important because it is the intermediary step between information encoding and decoding. Linguistic information cannot be learned if there is no place or way for it to be stored, and if it cannot be learned then it cannot be recalled for either input comprehension or output production. Finally, understanding how stored language information is accessed is important because that is where language production starts. Is language information accessed in a serial manner? Categorially? Is it interconnected? Is there one single, autonomous stream of processing that moves from basic units to integrated information, or is there a dual stream that includes top-down processing with contextual effects?

In addition to these three main questions, there are a number of problems language processing models must address. The first problem is that the division of linguistic information units is actually somewhat arbitrary. Phonemes do not easily separate in speech, as a speech stream is a constant flow of sound with varied amounts and types of turbulence, so phonemes bleed into each other (Goldrick et al., 2016). For example, the airstream in the word "let" is continuous from the lateral approximate /l/ through the open mid-front vowel /ɛ/ until the alveolar stop /t/ and so there is no distinct, measurable separation between phonemes. The same is true with word separation in speech—oftentimes the word boundaries in speech do not have distinctive pauses, or at least they are not as distinctive or organized as the spaces between words in written language are. The airstream in "let us sing" is only partially obstructed by the alveolar tap /ɾ/, the unvoiced alveolar fricative /s/, and the voiced velar nasal /ŋ/, and though these obstruents affect the airflow's turbulence, they do not make distinct, measurable separations that indicate exactly where each word begins and ends. Dividing language units by semantic or syntactic roles could either be overwhelming with the amount of information contained in each unit or result in overgeneralized categorical distinctions. These two outcomes reflect how even small morphological variations can vastly change a unit's semantic or syntactic categorization. Regardless of model type or foundational theories, any language-processing model should sufficiently address the three basic steps in language processing (encoding, storage, and access) as well as be able to account for occurring phenomena that seem to contradict or complicate our traditional notions of linguistic "units." An outline format provides a clean approach to understand the problems models must address when attempting to model natural human language provides.

List of Problems Models Need to Address:

1. Universality
    a. Can the model be applied to all languages?
    b. The assumption under this requirement is that a universal grammar exists because the faculty for language is biological.
        i. The approach linguistics use is called "generative grammar," which is a formative model of a native speaker's linguistic competence that allows linguistic production and comprehension.

2. Integration
    a. How do we connect mental representations (such as phoneme representations) with the physical aspect of language (stimuli sensation/motor movements)?
    b. How can a single, meaningful cohesive linguistic expression be properly produced as output from individual, separated functional units? How can the many different and separated units of information stimuli then be integrated into a single, meaningful perceived experience?
    c. Models need to account for context effects, for example:
        i. Retroactive repeated phoneme effect
        ii. Stress pre-entry effect
3. Phonemic ordering
    a. How do we differentiate between words that share the same set of phonemes but have different orders of the phonemes within the set (eg. “tap” and “pat”; “melon” and “lemon”)?
    b. How do we properly store, organize, and recall specific phonemic sequences in their precise and accurate order, especially since there is almost an indefinite amount of possible phonemic combinations?
4. Timing
    a. Do the theorized steps of a process in a model occur within the known and measurable time window we observe?

i. Even if a model can reproduce an output, it may not be able to do so in real-time.

5. Error Modeling

   a. Models cannot just account for normative language processing but also any types of language errors that can occur since errors are also governed by rules and constraints.

      i. This needs to be done without over-generating errors or modeling impossible error types.

      ii. The model needs to account for the probability that an error may occur, not for the certainty that an error will occur. This means that the model must account for error *likelihood* rather than simply showing that condition(s) X (Y, and Z) *always* lead(s) to an error.

   b. Models need to specify precisely at what linguistic level within the grammar an error can occur.

   c. Models need to be able to reproduce language production where there can be an intended phoneme/word *within* the middle of errored production, rather than the whole section be errored.

      i. Example: "rice **and** mats" instead of "mice **and** rats"

   d. Errors are often not simply one-for-one mismatches, so models need to be able to produce these gradated errors across ambiguous categorical boundaries.

6. Neurobiological Support
    a. Several cognitive models, models that many traditional linguists accept, are able to address the majority of the previous problems, but they do not take into account the structure and functionality of the brain.
    b. Models need to have neurobiological form, accuracy, and adequacy because language is a biological process that occurs for a significant part in the ***brain***.

Over the last several decades, linguists have tried to develop different language theories and models able to account for these challenges with varying degrees of success.

Chapter 3: Current Theories of Normal Language Processing

A number of language processing theories have been proposed with the area of focus and perspectives given dependent on which linguistic tradition the theory comes from. Each attempts to address the number of inherent problems when trying to make such implicit and intricate processes explicit. Though most of these models do not even mention spoonerisms, their validity or insufficiency can be exposed by testing whether or not they can adequately describe and explain the mechanisms that cause spoonerisms to follow the rules and constraints that make them occur.

There are currently two basic divisions of language processing theories: modular theories and connectionist theories. Modular theories break down each component of language processing into discrete, autonomous steps that become progressively integrated as the linguistic components in each step are built up through the cognitive system. These theories concur with the theory of localized neural functions where each cognitive function is associated with a specific neuronal structure. Connectionist theories, on the other hand, argue that language is functionally distributed throughout the brain as language processing itself is an interconnected network of systems. Language as a whole is predominantly lateralized to the left cerebral hemisphere (as with Broca's area and Wernicke's area), and there are a number of cognitive functions that appear to be located in specific areas of the brain, such as primary motor processing (used in speech) and sensory processing (used in listening). However, a significant amount of language processing—particularly complex processing tasks—occurs bilaterally with distributed connections across several specialized neural networks. Both modular theories and connectionist theories attempt to explain how specialized functions synthesize together to build a single integrated perception or action.

Modular Theories: Motor Theory of Speech Perception

One classical type of modular language production theory is the Motor Theory of Speech Perception (MTSP). In this theory, "modules" are functionally domain-specific processing devices with set, hardwired operations that reflexively react to highly specific input conditions (Fodor, 1983). Modules are designed for specific information processing tasks, such as syntactic parsing or phoneme recognition through feature detection, and modules work together as a system to support both the encoding of language during acquisition and production as well as the decoding of linguistic input. This theory accounts for the proposal of different "editing" mechanisms, such as the hypothesized lexical output editing mechanism (Baars, Motley, & MacKay, 1975) and the hypothesized semantic and phonotactic editing mechanism in the prearticulatory phase of speech encoding (Motley & Baars, 1976). Failures in components of these editing mechanisms could account for the production of verbal errors that follow some rules and constraints, just as spoonerisms do.

Wickelgran (1969) offers several serial order theories for encoding information that fall under the MTSP, as well as delineating each theory's potential flaws. He begins by presenting the assumptions psychologist Karl Spencer Lashley makes in his rejection of associative-chain theories of serial order[1]. Wickelgran explains that Lashley first assumes the existence of noncreative behavior (repeated behavior that occurs in the same

[1] Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior,* New York: Wiley.

manner each time), an assumption that could be used to describe elementary skilled motor movements observed such as grasping at an object or jumping or throwing a ball. This noncreative behaviour would assumedly be controlled by a single sequence of internal representatives for each of the elementary motor responses (EMRs) at the central articulatory level. For example, the internal representatives of each phoneme in the pronunciation of the word "right" (which could be /r/, /ī/, /t/) would sequentially activate the phoneme that follow it (so /r/ would activate /ī/ and /ī/ would then activate /t/). The third assumption Lashley makes according to Wickelgren is that there is a finite set of equivalence classes on an infinite set of response sequences, meaning that the internal representations of all EMRs would be identical regardless of the contexts around the specific EMRs.

Based on these assumptions, Lashley proposes behavior sequences are produced by "context-free coding of elementary motor responses." The context-free coding of EMRs argues that words are coded as sequences of phonemes in the speech system without pairwise associations between phonemic representations. Context-free coding of EMRs contrasts with the theorized context-free associative system that argues the internal representatives of EMRs are associated to articulatory representatives. This means that, according to the context-free associative system, when a word is pronounced the appropriate phonemic representatives are selected from the unordered set of phonemic representatives for the word based on the association strength to the word representative. One problem with the associative theory that context-free coding of EMRs avoids is with words that share phonemes in different orders (such as with "melon" and "lemon") because the unordered sets of phonemes are identical—if the basis for the serial ordering of the phonemes is phoneme-to-phoneme associations, the phonemic order could not be rearranged to produce both words. Additional support against the associative theory is that there is not enough time for the first phoneme's auditory and/or kinesthetic feedback to trigger the next phoneme, as well as that the pronunciation of a phoneme is influenced by the phonemes that follow.

Wickelgran agrees with Lashley that context-free associative memory is inadequate because it fails to address the ability to differentiate between words that share the same set of phonemes with different orders as well as its inability to model language processing in real time, so he proposes and analyzes a number of alternative theories on serial order that may explain how phonemes are ordered to produce speech on the word level. The first alternate theory Wickelgran analyzes is "context-sensitive associative memory," where the internal representatives of each unit in different local bilateral contexts are different (such as the phoneme /i/ in the word "sit" is represented as ${}_{s}i_{t}$, whereas /i/ in the word "hip" is represented as ${}_{h}i_{p}$ rather than simply represented as just *i*). The difference in internal representations with respects to context allows order by priming each phonological unit to prevent confusion by individual and by distinct encoding of allophones with distinctive features (like stress) included. While context-free associative memory could account for basic serial order problems for noncreative behavior sequences of one identical pair of EMRs, this theory does not account for the ability to pronounce identical pairs of phonemes followed by different phonemes, as in the case of /lampblack/ (where two pairs of /la/ are followed by the two different phonemes /m/ and /k/).

A second alternative theory is "contingent associative memory," which claims that there is a strong interaction between a word representative as a whole and the associations between phoneme representations as each word is coded as a sort of "pre-ordered" set of phonemes. This theory allows room for spoonerisms to occur because, as Wickelgran explains, the transposition of phonemes occur with higher probability in connection with words or phrases having repeated phonemes than with words or phrases that do not have any repeated phonemes. Relatedly, the next theory Wickelgran offers is the "multiple associative memory," stating that a word is coded as an ordered set of phonemes, which is similar to the contingent association theory except that each phoneme representative in each word is different from each other. This would mean that each word representative would be made up of a set of unique phoneme representatives (such as that /i/ in "sit" has one phoneme representative and the /i/ in "hip" has another, separate phoneme representative).

The final alternate theory that Wickelgran analyzes is the "nonassociative memory" theory that claims locations are used to order the encodings of any member of the relevant set of internal representations. This would require there to be at least as many locations in a nonassociative buffer store as there are phonemes in the longest word or phrase that conceptually forms a single unit since each phoneme representation would be "stored" in each location. The problem with this theory is that if memory were nonassociative, errors with repeated elements in spontaneous speech (like spoonerisms) may not be possible.

Wickelgran asserts that all four of these alternate theories of serial order in noncreative behavior (context-sensitive associative theory, contingent associative theory, multiple associative theory, and nonassociative theory) can be conceptualized as one theoretical continuum of associative activation as each theory shares some similarities with the next. To determine which of these four theories could account for human speech production, Wickelgran tested these theories in the context of repeated-item phenomena, the coarticulation effect, and the pronunciation of full phrases. Wickelgran concludes that the context-sensitive associative theory may be the most likely theory to account for the ordering of phonemes in natural and spontaneous speech because it handles syllable structures and distribution the most effectively.

A year later, MacKay (1970) would examine some of the theories Wickelgran discusses in his own analysis of spoonerisms produced by German speakers in natural speech. This analysis found a number of patterns: the identical phonemes typically either precede or follow the reversed phonemes, repeated phonemes that follow a spoonerism are more frequent than repeated phonemes preceding spoonerisms, and the reversed phonemes typically had similar articulatory form and syllabic position. MacKay then compared the data with a number of previously proposed theories: chain association theory, the similarity hypothesis of phonemes, the proximity hypothesis, the syllabic similarity hypothesis, the syllabic structure hypothesis, the Relational Memory Theory, and the linguistic universal hypothesis. He found first that the results of his analysis contradict the chain association theory—the theory predicts that repeated phonemes precede spoonerized phonemes more frequently than repeated phonemes following spoonerized phonemes because the chain of associative bonds are unidirectional, but the data demonstrates that repeated phonemes follow spoonerized phonemes just as frequently, if not more frequently than the preceding phonemes. The similarity hypothesis that states phonemes are more frequently switched if they share similar articulatory characteristics is supported by the data in terms of openness, voicing, and nasal characteristics (the spoonerized phonemes had these characteristics in common), but spoonerized phonemes did not frequently share place of articulation (since front and back consonants switched more frequently than the consonants with closer place of articulation).

The proximity hypothesis is supported with this data because as phonemes were in closer proximity (both within and between words), they were more frequently spoonerized (MacKay, 1970). The syllabic similarity hypothesis (in which spoonerized phonemes are in the same syllabic position) is supported by the data for both consonants and vocalic reversals. The syllabic structure hypothesis theorizes that a specific syllabic position is more likely to be spoonerized than other syllabic positions. The data also supports this theory because most spoonerisms occurred specifically in the initial syllabic position. This may be because vowel + consonant$_{(final)}$ /consonant cluster$_{(final)}$ both form subgroups that resist being broken up, making the syllable-initial the easiest to spoonerize. Word-initial phonemes spoonerize more frequently than non-word-initial phonemes, possibly supporting Relational Memory Theory, which suggests that the collection of a word’s phonemes are stored in abstract relational forms and that the first phoneme of a word is separately stored (which would make the word-initial phonemes easier to switch if they are separated from the rest of the word). The linguistic universal hypothesis, claiming that there is a universal underlying language mechanism that spoonerisms may reflect a part of, seems to be supported because the phoneme repetition effect is independent of language type since the effect is seen in at least Latin, French, Greek, Croatian, and German.

After comparing his analysis of the data with these different theories of speech serial order, MacKay explains that though the context-sensitive chain association theory seems to explain the serial order of speech, the theory as it is does not explain the retroactive repeated phoneme effect, the stress pre-entry phenomenon, the effects of syllabic position on spoonerisms, the phonetic similarity of the reversed phonemes, or how phonemes intervening between reversed phonemes are produced without error (an effect that actually closely parallels visual illusions and the correctly perceived forms intervening between visual stimuli).

Modular theory as a whole has a number of advantages. For one, it provides a basis for an editing mechanism, the failure of which would account for verbal slips and other types of language production errors. However, modular theory has a difficulty explaining how a cohesive linguistic expression can be properly produced from single, separated functional units. Even some of the most supported modular theories like the context-sensitive chain association theory are unable to account for the occurrence of several phonemic phenomena or errors.

Connectionist Theories: Cohort and TRACE

Cohort Model

The second division of language processing theories falls into the category of connectionist theories. One type of these connectionist theoretical speech production models is the "Cohort Model," first proposed by Marslen-Wilson and Welsh in 1978 under the label "active direct access model" in an attempt to develop an interactive model of spoken word recognition that parallels bottom-up information processing with the contextual constraints affecting the process's outcome (Marslen-Wilson, 1987). The cohort model began as a modular model, but over time it became a connectionist model of lexical access with the purpose of modeling how words are recognized and retrieved from the mental lexicon through a serialized selection process.

In the cohort model, the process of spoken word-recognition is segmented into three basic steps: **access**, in which the speech input (the physical acoustic sound stimulus) is mapped onto the lexical form representations, **selection** of the best-fitting match of the word-form representation on the lexical map to the speech input, and **integration** of the semantic and syntactic information with the selected lexical form onto higher level processes (Marslen-Wilson, 1987). Each lexical form representation entry in the mental lexicon is thought to correspond to one discrete computationally active recognition unit. Each recognition unit represents a functional connection between the acoustic-phonetic information and contextual (syntactic and semantic) information that belongs to the lexical entry.

One question the cohort model attempts to answer is whether speech perception processing is solely a bottom-up serial process or if it is a dual-stream information process with top-down information processing interacting with and influencing speech perception in addition to the bottom-up information processing. Cohort gating experiments do indicate that feature extraction and structure building are involved in speech perception, suggesting that there is an integration of speech signals and semantic representations rather than a compartmentalized modular-type process.

In the revised version of the cohort model, Marslen-Wilson (1987) asserted that the first step in speech perception (accessing the mental lexicon) is solely a bottom-up process. In this model, the acoustic sound is the only stimuli being processed without any other information influencing mental lexicon access. Then the system moves on to the selection phase where contextual constraints begin to factor in and affect the process's outcome. He argues that these systems work in parallel where different information sources (phonetic, semantic, and syntactic) eventually integrate together to synthesize the final perceived output, though these paralleling systems never actually interact or influence each other but rather work autonomously in either the form-based access or form-based selection steps.

Though there is general consensus among researchers that there is significant dual processing at all stages of perception, not all connectionist model researchers agree that the two parallel streams of information never interact. In fact, the research by Goldrick et al. (2016) brings the connectionist principles of the cohort model for speech perception into a framework for speech production and demonstrates how the dual streams do interact and affect each other. In their research, Goldrick et al. looked at phonetic-based verbal errors elicited through tongue-twisters. They used a novel algorithm to detect and locate linguistically relevant acoustic properties in speech samples, which allowed large datasets to be analyzed for speech errors and error types in a more accurate and unbiased manner with greater reliability than human coders. What they found was that speech errors are partially a reflection of the intended sound, exemplifying how phonetic representations are cognitively categorized by a gradation of sound representations when in reality sound waves exist on a continuum of frequency variation. Since articulated sounds exhibit these slight variations along the graded scale, each utterance of one word or sound will not be exactly the same even though they are understood to fall into the same phonemic category that represents the word. These variations in articulation are even more explicit in the context of tongue twisters: Goldrick et al. found that verbal errors are not simply complete one-for-one sound substitutions but rather a form of phonetic blending, in which "ααα.βββ." does not completely switch to "βββ.ααα" but to "ββα.αβα." In this process, the phonetic representation of each switched phoneme (the categorized sounds of /α/ and /β/) exists on an articulatory gradation between /α/ and /β/. This means that though the phonemes are perceived as having completely metathesized from /α/ to /β/ and from /β/ to /α/, the phonetic characteristics of "ββα." exists somewhere between "βββ." and "ββα.," and likewise for "αβα." As such, it is possible that spoonerisms may also occur as a result of confusion between the existing phonetic variability and so also are characterized by a mixing along a gradient rather than just substituting one unit for another.

Goldrick et al. (2016) identified two types of cognitive processes that underlie speech production: the planning processes and the articulatory processes. As the authors discuss, the planning process involves cognitively constructing an abstract specification of the articulation targets (which are the abstract ideas of which proper articulator movements must be performed for proper pronunciation). The articulatory process involved identifying the exact real motor movements the articulators need to follow to execute the plan. The importance of these authors' findings is that articulators may have an effect on the pronunciation of words, which may in turn affect the presence or absence of a spoonerism. This effect demonstrates how the cognitive top-down processes and the bottom-up motor articulatory processes influence and interfere with one another to produce phonetic blending.

TRACE

While the cohort model had started out as a modular model and eventually became a connectionist model, the TRACE model first began as a localized connectionist model of speech perception and evolved into a more distributed—but still connectionist—model. The purpose of TRACE is to simulate the process taken for identifying lexical effect on comprehension and retrieval (McClelland & Elman, 1986). According to TRACE, when we retrieve stored lexical items, they are retrieved in competition with each other along with inhibiting units so that the most competitive (and therefore the best candidate for correct selection) should win and be selected. However, sometimes the emergent winner is not the best, correct choice. This can be accounted for if top-down information flow activated by lexical entries overrides the bottom-up information processing of phonetic representations.

There are actually two TRACE models: TRACE I models the phonological processing and pre-lexical effects in speech perception, and TRACE II models word recognition and lexical influences on phoneme recognition. TRACE's advantages are that it can model cohort effects while simultaneously modeling possible top-down influences that can arise from neighboring cohorts, and it can model effects like coarticulation effects and categorical perception through lateral inhibition.

Summary

Unfortunately, TRACE and the other models discussed in this section are cognitive models based on computer processing. These models do have their merits as they have provided a platform for examining language processing and have provided research information that can be used and analyzed. The modular theories, in their focus on information ordering, account for the proposed editing mechanisms and begin to account for possible contextual effects. Yet modular theories face problems in the proposal for how information is ordered and integrated when presented with spoonerisms because, as Goldrick et al. (2016) shows, language is not so easily segmentable with clear boundaries. If there are no clear and discrete boundaries, self-contained modules defined by distinct boundaries could not exist. On the other hand, the connectionist theories focus on integration and contextual effects as well as ordering. They are able to model word recognition through the serialized selection process from the mental lexicon, showing how to move from the physical stimuli to the integration of contextual information. Still, the connectionist theories are based on computational processing, remaining reliant on distinct linguistic feature segmentation and so are unable to process natural speech streams (which are produced without clear segmentation boundaries).

Indeed, none of these models take neurobiology into account. Even if these models were to be able to sufficiently account for the universality of language, for integration of information, for contextual effects, for the storage and access and ordering of information on each linguistic level, for timing, and for how language errors occur, they would be ignoring our biological nature. These models may be able to successfully predict outputs, but computational prediction does not equal neurological (or even computational) certainty, and so at best these models can only be analogical to the processes actually occurring. We are biological beings and language processing happens in the brain rather than a computer. If a process cannot happen in the brain, it cannot be modeling natural human language. An accurate language model needs to be built on a neurobiological foundation so that it can show how the brain seamlessly processes and integrates information as it moves from the physical materiality of language in its unsegmented state to the cognitive categorical perception of language. One way to approach the task of building a language model with neurobiological validity is by neurobiologically examining speech production errors like spoonerisms.

Chapter 4: Linguistic Description of Spoonerisms

Most theoretical language processing models, regardless of their foundational approach, all derive their fundamental premises and methods from generative grammar's framework for analyzing languages. Generative grammar is a linguistic theory first proposed by Noam Chomsky in *Syntactic Structures* (1957) and has since been revised by Chomsky and other linguists. The essential motivation behind generative grammar's development was and is an attempt to understand how children can naturally acquire something so complicated without needing explicit explanation or direction from adults. The theory first makes a few assumptions about language. First, generative grammar assumes that language is a biological process, and therefore is an innate function rather than a learned behavior. It also assumes that all natural human languages are comprised of a shared set of finite principles, within which are a finite set of shared parameters that may be optimally set to one of two settings for each principle. It is through various setting combinations of these parameters that linguistic variations exist across languages. The main conclusion of generative grammar is that based on the assumptions that language is a biological process and that all languages have some variation of parametrical settings within the same finite principles, language—and therefore grammar—is universal.

The goal of generative grammar is to model a native speaker's implicit competence through a series of ordered rules that can produce a linguistic output along with filters and constraints that then limit the linguistic output. This series of rules, filters, and constraints should define the outermost limit of linguistically well-formed structures in a language so that the model produces all and any natural utterance that can be spoken by a native speaker (therefore is “generative”) without ever generating a structure deemed unacceptable to a native speaker (also known as “over-generating”). If the grammar can successfully generate all and only well-formed structures according to the intuitive judgments of native speakers, it would successfully and explicitly demonstrate the speaker’s implicit knowledge, allowing linguists to study and understand the intricate and naturally-hidden system behind language processing.

In developing a grammar, linguists begin at the observational level by looking for patterns in linguistic structures. Once the data have been exhausted, the observations can be analyzed to form the descriptive level. The ultimate goal of a generative grammar is to achieve explanatory competency, in which the fundamental structure of a language is completely explicit and is capable of successfully predicting any possible grammatical linguistic occurrences without over-generating. To develop a generative grammar in the context of spoonerisms, the rules would describe the linguistic conditions in which spoonerisms can occur, and the constraints would describe the conditions in which spoonerisms cannot occur. This way, a grammar will demonstrate competency when its rules describe any and all occurrences possible but will never predict and describe any spoonerism that does not and cannot occur. Such grammatical description assumes that every linguistic process is rule-governed, even language errors like spoonerisms. While spoonerisms seem to be simply a type of verbal error, spoonerisms actually follow certain rules or required conditions for where they will or will not occur. In the past few decades, linguists and psycholinguists have conducted studies and published work that focus on such contexts that seem to influence the likelihood of spoonerism production in an attempt to isolate and identify the specific conditions in which spoonerisms can occur. Though they did not propose any concrete rules or constraints, in their studies they observed that word position, phoneme position and quality, lexical validity, and semantics may at least facilitate spoonerism production. Based on the description of the linguistic conditions in which spoonerisms occur, we can begin to build a grammar that attempts to determine the underlying rules and constraints that lead to spoonerism production.

Spoonerism Observation

Linguists have identified several environments in which spoonerisms occur, providing a basis for observational competency. These environments can be described on various levels:

1. **Morphologically**

A study of the morphemes (the smallest grammatical units of a language like root words and affixes) involved in spoonerism production shows that most spoonerized word pairs tend to switch sound units occupying the same position in each morpheme—if the sound unit of the first word occurs in the initial position, then it is more likely to switch with the sound unit occupying the second word's initial position rather than a middle sound unit or a sound unit in the end position (Motley, 1973). For example, with the word pair "top hat," the initial sound unit "t" most frequently switches with the initial sound unit "h" rather than "a" (middle) or "t" (end). Thus, the spoonerism will most likely become "**h**op **t**at":

Top **H**at → **H**op **T**at

This increased frequency seems to also be true for middle position-to-middle position and end position-to-end position—in these instances, "top hat" could become "t**a**p h**o**t" for middle position or "to**t** ha**p**" for end position:

T**o**p H**a**t → T**a**p H**o**t   To**p** Ha**t** → To**t** Ha**p**

However, sound units from two different positions do not often switch, like "to**h** **p**at" (an end position-to-initial position switch):

To**p** **H**at → *To**h** **P**at[2]

Out of the three, however, word-initial position switching seems to occur much more frequently than middle position or end position switching. These variations in frequencies suggest that the switching sound units are primarily influenced by the similarity of word position and secondarily by the sound units' position themselves (word-initial or not).

2. **Phonetically**

The phonetic study of spoonerisms examines the acoustic and articulatory processes and components of speech sound occurring (or unexpectedly not occurring) during the physical production of spoonerisms.

3. **Phonemically**

A study of the phonemes (the smallest meaningful units of sound in a language) involved in spoonerism production shows that the phonemes being switched in spoonerisms retain their phonological categorical integrity (Motley, 1973). This means that a phoneme will

[2] An * indicates that the example is ungrammatical

never change to a different phoneme when switched. For example, if the final phonemes in “Yard Barn” switch, the /d/ and /n/ will stay a /d/ and /n/ in their new location rather than changing to a different sound like /t/ or /g/:

Yar**d** Bar**n** → Yar**n** Bar**d** / *Yar**g** Bar**d** / *Yar**n** Bar**t**

4. Phonotactically

Phonotactics focuses the study on the meaningful restrictions on phoneme distribution within syllables that affect syllable building, in this case the restrictions on intrasyllabic phoneme distribution in spoonerisms. A distribution restriction on how phonemes can be sequenced to form a syllable is called a “phonotactic constraint.”

Straddling the phonetic and phonotactic study of spoonerisms is the concern for place of articulation. The sound unit’s phonemic value or phonemic environment seems like it should have an affect switching frequency. Indeed, one analysis by Motley (1973) found voiced bilabial nasal stops (/m/), voiced bilabial stops (/b/), unvoiced bilabial stops (/p/), voiced alveolar approximate (/r/), alveolar lateral liquid approximate (/l/), bilabial glide (/w/), unvoiced velar stop (/k/), and high front tense vowel (/i/) to be the most frequently switched phonemes, suggesting that consonants are more likely to be switched than vocalics. Another study stated that the in the analysis of over a hundred spoonerisms produced in natural German speech, reversed phonemes usually had similar articulatory form with respects to voicing, nasality, openness, and syllabic position (MacKay, 1970). Yet, the difference between specific phonemes with a greater tendency to spoonerize from other specific phonemes was *not* statistically significant (Motley, 1973) and the spoonerized phonemes’ place of articulation is more frequently different than would be expected from chance (MacKay, 1970), indicating that the idea that certain phonemes are likely to have a greater spoonerism frequency is not supported.

The observation that the place of articulation between the spoonerizing phonemes is significantly different also suggests that there is some sort of physical factor influencing spoonerism production. Surprisingly, most language models focus solely on the phonological grammar but fail to examine the physical motor action involved or the physiological effects on language, and so this observation points to an insufficiency in the existing models.

5. Lexically

Lexical studies of spoonerisms involve studying whether or not spoonerisms are affected by the lexical validity of the context (where the surrounding words exist in the language's vocabulary) or the lexical validity of the word pair itself (pre- and post-spoonerized form). The lexical status of both the targeted word pair and the context around the word pair also constrain spoonerism occurrence (Baars, Motley, & MacKay, 1975). A "lexically valid" word pair consists of two words that exist in a language's lexicon, also known as a language's vocabulary. One study examining the output editing for the lexical status of spoonerisms found that lexically valid spoonerism outcomes occur more frequently than lexically invalid (nonsense) spoonerisms in a lexical context (Baars et al., 1975).
For example, when

Top **M**ap → **M**op **T**ap

"mop" and "tap" both exist in English's lexicon and so are considered lexically valid, compared to the lexically invalid "gad" and "boof" outcome with

Bad **G**oof → ***G**ad **B**oof

A “lexical context” in this study involved presenting the target word pair surrounded by other lexically valid word pairs first, such as if “Could Gore” and “Cook Goes” preceeded Good Cot, so as to prime participants to spoonerize

Good **C**ot → **C**ould **G**ot

since they were expecting to read and say real, lexically valid words. On the other hand, a “nonsense context” in this study involved presenting the target word pair as words surrounded by nonsense words so to prime participants to spoonerize targeted word pair into a lexically valid word pair, such as “Setch Rab” and “Sul Ret” preceding “Rafe Sode”

Rafe **S**ode → **S**afe **R**ode

or to prime participants to spoonerize the targeted word pair into a lexically invalid word pair

Rabe **S**ofe → **S**abe **R**ofe

However, this study also found that there was no general tendency for the error rate to favor lexical outcomes over nonsense outcomes unless there was a reason for the participants to expect real words, meaning that participants produced significantly more lexically valid spoonerisms when they listened to the targeted word pairs in a lexical context compared to those produced in a nonsense context.

6. Semantically

Semantic studies of spoonerisms—the study of how meaning (meaning on both the word and context level) interacts with spoonerism production—show that semantic conditions also affect the frequency of spoonerisms. A study conducted by Motley and

Baars (1976) found that the targeted word pairs which were semantically synonymous to the preceding word pairs were significantly more likely to spoonerize than the targeted word pairs unassociated semantically with the preceding word pairs. For example, “Good Group” and “Pleasant People” preceding “Mice Knob” would cause

Mice **Kn**ob → **N**ice **M**ob

or “Ill Bishop” and “Stricken Priest” preceding “Pick Soap” would cause

Pick **S**oap → **S**ick **P**ope

They also proposed that there may be some form of semantic editing process in the prearticulatory phase of speech encoding in addition to the previously proposed phonotactic and lexical editing processes. From this, Motley and Baars hypothesized that the frequency of spoonerisms for targeted word pairs preceded by both semantic and phonological interference would be significantly greater than the frequency of spoonerisms for targets preceded by phonological interference only. They found that the speech encoding systems of the participants were sensitive to semantic influences as participants responded to semantic priming related to the spoonerized form of the word pairs but not the targeted (unspoonerized) word pairs, suggesting that the semantic priming linked to spoonerized forms of word pairs influenced participants to become biased to the spoonerized form.

Rules and Constraints

The levels of observations described above can be organized as a framework of rules and constraints necessary for spoonerism production.

Rules

Spoonerism Rule: Switch some part of Word1 with some part of Word2

AB**C** XY**Z**→ AB**Z** XY**C**

While a spoonerism is a type of metathesis where some part of one word switches with some part of another word, like "**W**all **H**ole" becoming "**H**all **W**ole," this rule over generates because "*Ca**oons** and Racc**ts**" (from "Ca**ts** and Racc**oons)**" or "***Nai**e **To**ls" (from "**To**e **Nai**ls") do not occur and so can be considered "ill-formed" or ungrammatical.

Morphemic Rule: Move like position to like position on the morphemic level

ABC **X**YZ → **X**BC **A**YZ

This produces a distributional restriction that accounts for spoonerisms on morphemic level. The morphemic rule, however, is insufficient as well because not all morphemes can be switched, even if they are in the same position within the word:

Un**intention**ally Re**member**ed → *Un**member**ally Re**intention**aled

Another reason the morphemic rule is not sufficient is because it does not take into account the patterns of phonemic distributions that we see, as well as those we would expect but don't see. Though the phonemic value of the sound unit and the phonemic environment around the units seem to be important linguistic conditions for eliciting or preventing spoonerism production, phonemes themselves do not appear to affect switching likelihood (Motley, 1973).

Phonologic Preservation Rule: Preserve the phonological category for any phoneme being switched during metathesis:

/α/→ /α/

***/α/ → /β/**

Practically, this would mean the phoneme switched will never change into a different phoneme

/p/→/p/ and /p/ /k/ → /k/ /p/

While

***/p/ → /mp/ */p/→/k/**

In following a language's phonological grammar, the phonemes in a spoonerism must retain their phonological integrity and must preexist in the language's phonological inventory. This means that the phonemes that switch are either completely X phoneme or not X phoneme at all, never half X phoneme (Motley, 1973).

Phonotactic Constraint Rule: Any metathesized phoneme must obey phonotactic constraints of native language

Squirrel **Ch**asing→ **Ch**uirrel **Sq**uasing

Sleepy **Sh**eep → ***Shl**eepy **S**eep

In this example, "*Shleepy Seep" is considered ungrammatical in English because the onset (the beginning of the first syllable) has too many consonants before the vowel. In English syllable building, there is a constraint against alveopalatals preceding lateral approximates in a onset consonant cluster. On the other hand, "Chuirrel Squasing" is grammatical in English because it exhibits phonotactic preservation and adheres to

English's phonotactic constraints. Another example of a phonotactic constraint rule violation can be seen in

[p^hIt. kɝɹn] → *[kIt. p^hɝɹn]

In English, when there is a stop in an onset position before main stress, it has to aspirate. If two plosives—one aspirated and one unaspirated—metathesize, the aspiration has to be decoupled from the first plosive and remain in the onset position preceding main stress. In the example above, the aspiration follows the metathesized plosive (/p/) to the onset of an unstressed syllable and the plosive /k/ remains unaspirated despite preceding the syllable with main stress, thus breaking this English-language phonotactic constraint.

Constraints

These rules are good because they produce spoonerisms but they over-generate, so we need to introduce a few probabilistic constraints. There are three constraints that can be extrapolated from the observations: sonority optimization, lexical fit, and semantic fit.

Sonority Optimization Constraint: optimize sonority in a metathesized sequence

Sonority is the least turbulent air flow, associated with vocality of a sound. Phonemes can be ordered on a sonority scale (Table 1).

Table 1. ***Sonority scale of phonemes from least sonorous (voiceless oral stops) to most sonorous (low vowels).***

Oral Stop		Fricative		Nasal	Liquids	Semivowels	High Vowels	Low Vowels
Voiceless	Voiced	Voiceless	Voiced					
p, t, k	b, d, g	f, θ, s	v, ð, z	m, n, ŋ	l, r	j, w	i, u	a, ɑ

A number of factors contribute to sonority constraints, such as intensity of the sound waves, the temporal compression or expansion of a sound unit, and formant transitions, all of which can vary depending on context (Finely, 2017). Sonority as a process is important because it is one way we perceive boundaries between syllables, which in turn helps us process linguistic sound input more efficiently. The sonority sequencing principle states that the nucleus (center) of the syllable is the most sonorous (vowel-like) part of the syllable (meaning it has a voiced and relatively unobstructed vocoid). The syllable structure generally builds from the least sonorous phoneme at the beginning of the syllable to the most sonorous phoneme at the nucleus before the syllable ends in a less sonorous phoneme post-nucleus. The constraint, when active, optimizes the sequencing of sounds in a syllable to fit the sonority sequencing principle structure. This process occurs because optimized sonority accelerates the cognitive processing of the syllable because it allows syllabic boundary identification to occur much more quickly.

For example,

Nobel **L**aureate → **L**obel **N**auriet

This spoonerism in this example optimizes sonority through the metathesis because it separates the geminate (the double "l" between Nobe**l** and **L**aureate) to clarify the syllabic boundary in a sequence of approximates. Moreover, this decoupling of the liquid geminate makes the sonority symmetrical within the word between the two syllables in "Nobel," which makes it even easier to process.

Another way to optimized sonority is to switch the assignment of stress. For example, In Donald MacKay's experiment on motor stress pre-entry, participants were asked to rapidly repeat sequences of four syllables “tay,” “gay,” “bay,” and “day,” with prosodic stress on a certain syllable designated by the experimenter (1971). The study found that participants would accidentally switch the stressed syllable with the preceding syllable. This is important because stress can also indicate perceptual syllabic cues.

Lexical Fit Constraint: Spoonerisms must lexically fit the linguistic context

“The cat chased the **mice** and **rats**” → “The cat chased the **rice** and **mats**”

The lexical status of both the targeted word pair and the context around the word pair constrain spoonerism occurrence as lexically valid spoonerism outcomes occur more frequently than lexically invalid (nonsense) spoonerisms in a lexical context (Baars, Motley, & MacKay, 1975). Therefore, spoonerisms are constrained to form from a lexically valid word pair in a lexically valid context and metathesize into a lexically valid word pair.

Semantic Fit Constraint: Spoonerisms must semantically fit the linguistic context

Speech encoding systems are sensitive to semantic influences because when a word pair is exposed to semantic priming (in which the linguistic context around the word pair semantically relates more to the spoonerized form of the word pair than to the unspoonerized form), the speaker becomes biased to the spoonerized form (Motley & Baars, 1976).

Explanatory Adequacy

These rules seem to decently express the limits between grammatical and ungrammatical formations of spoonerisms. The presence of lexical and semantic contextual constraints makes sense because a number of language models have tried to determine the extent to which linguistic context affects language production. The transition from the structural rules pertaining to morphology, phonetics, phonology, and sonority to the context-sensitive rules mirrors the transition within language processing from structure to information integration. Context is considered to be a top-down process because it deals with non-basic information that may or may not affect perception or production of linguistic information, whereas "bottom-up" processing involves building up basic informational units like phonetics to form a cohesive meaningful perception. While the cohort model considers context to be a parallel but completely separate and noninteractive process (Marslen-Wilson, 1987), the grammar seems to indicate that spoonerism production is clearly just as much a top-down process as it is a bottom-up occurance.

From a more global perspective, these rules together demonstrate a commonality: the linguistic conditions on each level—the morphological, the phonemic, the syllabic, and the prosodic level—provide a cue for a spoonerism to occur. All of these processes involved in spoonerism production share an emphasized problem concerning how the brain segments physical streams of information into appropriate cognitive categories. Normally when we think about language, we see clear boundary markers between distinct units that are conjoined to build larger units that can then make even larger units, all to easily externally convey internal feelings and ideas. In writing for example, the smallest unit we see is a letter. Letters join together to make words, words join together to make phrases, phrases make sentences, sentences make paragraphs, and so on. Traditional language models are based on this unit segmentation, and so they require clear boundaries between categories. The problem is that in reality, language processing like speech does not actually have these clear and distinct categories (Figure 1).

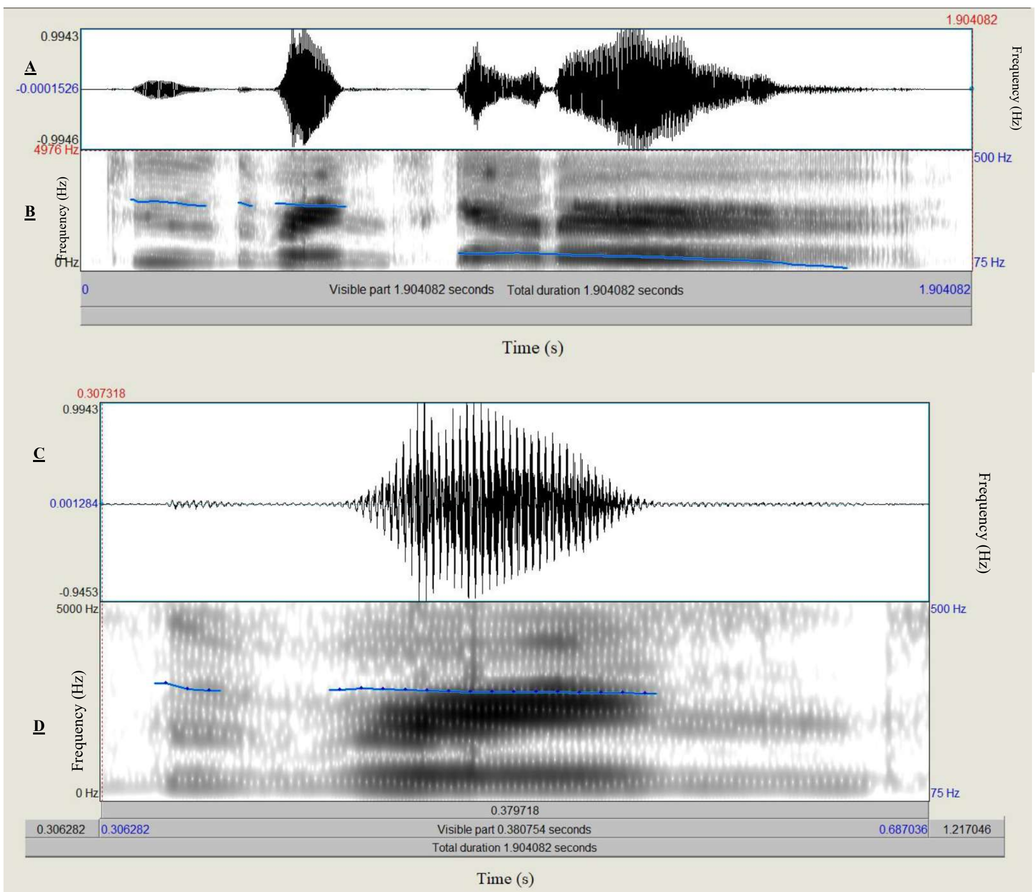


***Figure 1.* Spectrogram illustrating the phrase "To the brain scanner." Top graphs (A, C) show the wave form of the sound over time. Bottom broadband spectrograms (B, D) show the spectral energy of the sound over time. Blue lines in the broadband spectrograms represent the speaker's pitch. *A* and *B* show the full phrase, while *C* and *D* illustrate the transition from "the" to "brain" in the phrase "the brain." Recorded using PRAAT from http://www.fon.hum.uva.nl/praat/.**

The units we conceptualize and that traditional language models rely on do not exist—a speaker does not pause in between each syllable, each word, or even between every sentence. In terms of physics, the stream of sound waves is continuous as each phoneme influences the next and syllables bleed together even across sentence "boundaries" (Goldrick et al., 2016). Because of this lack of decisive segmentations, traditional language models cannot process live speech signals nor can they accurately model natural speech production (this inability is clearly seen in speech produced quality by voice recognition and simulation programs like Siri or Alexa).

Consequently, we can describe the distribution of features and identify each level of influence linguistically, but traditional language processing models and observational studies have to stop at the descriptive level of the generative grammar because they cannot move onto the explanatory level without getting into cognition and biological processing. First off, one study found that while phonemic categorization remains intact during metathesis, phonetic blending does in fact occur (Goldrick et al., 2016). In addition to the phonemic rules that spoonerisms follow, it seems there is a motor component to how phonetic features influence spoonerism production. If such a motor component does in fact exist, it would need to be accounted for, but the generative grammar framework of rules and constraints does not have a space to include motor control. Therefore, generative grammar by the very nature of it traditional framework can only inadequately describe language processing at the observational level.

Evidence for this motor component includes a study by Goldrick et al. (2016), which explains how articulators may have an effect on the pronunciation of words when there is confusion between existing phonetic variability, which may cause mixing units along a gradient. Their research began by looking at phonetic-based verbal slips elicited through tongue-twisters and using an algorithm to detect and locate linguistically relevant acoustic properties in speech samples. Goldrick et al. explained that if the mechanism for the planning process for selecting the appropriate phonemic units is disrupted but the intended target's specifications remain partially active, the error produced will be distorted towards the intended target, producing an articulation that combines properties of both the originally intended target and the errored target.

They also explain how articulations are on a graded scale of sound, and how each utterance of one word or sound will never quite be the exact same as the other utterances. For example, one important aspect of phonemic distinction is the voice onset time (the time between when the release of airflow starts and when the vocal folds start vibrating). In English, voiceless sounds like /p/ tend to have relatively longer voice onset times, whereas voiced sounds like /b/ tend to have a shorter voice onset time. However, in natural speech these voice onset times can vary—if the voice onset times change too much, they may begin to sound like their voiced or unvoiced counterpart while retaining some of the original phoneme's elements. The results from Goldrick et al. (2016) support the hypothesis that speech errors in general involve the partial production of the intended sound unit along a grading of sound representations rather than the production of a sound unit in distinctive and separate sound categories.

This variation along the sound gradient is influenced even more when the utterance is in a context like that of tongue twisters, so that speech errors are not simply an item substitution of one sound unit from one category for another sound unit from another category but rather a phonetic blending. An abstract representation of this change could be that "ααα.βββ" does not become "βββ.ααα" but rather "ββα.αβα." For example, in the word pair "pet bet," the /p/ does not completely switches to /b/ and vice versa when the word pair is spoonerized into "bet pet"; rather, the first word's initial phoneme is produced by a primary activation of /b/ with a little activation of /p/, and the second word's initial phoneme is produced by primary activation of /p/ and a little activation of /b/. Goldrick et al.'s research indicates that the physical articulators may have an effect on the pronunciation of words, which in turn may affect the presence or absence of a speech error. As the authors discuss, the planning portion of speech production involves the construction of a relatively abstract specification of the articulation targets (the abstract ideas of proper articulator movements to perform for proper pronunciation), and the articulatory portion involves identifying the exact motor movements the articulators must follow to properly execute the plan. Following the suggestion that speech errors may *not* just be substituting one unit for another but a mixing along a gradient, it is possible that spoonerisms also occur as a result of confusion between the existing phonetic variability.

Another problem with attempting to formulate a generative grammar for spoonerisms based on these observations is their inability to account for cognitive processes, such as internal editing processes. In their research, a number of linguists have proposed the existence of an internal editing mechanism that is active during language processing, and that spoonerisms may be a result of a failure in this editing process

(Baars et al., 1975; Motley & Baars, 1976). Baars et al. (1975) explain that the ability for speakers to produce unintentional verbal slips (like spoonerisms) is due to a failure in the editing process. They suggest that the lexical editing phase could involve a mechanism that checks for the lexical status of an anticipated utterance. For example, "rats and mice" has better fit in the sentence "The cat chased the **rats and mice**" than the phrase "mats and rice" to make "The cat chased the **mats and rice**" because cats are often known for hunting rats and mice but not so much for their mat and rice hunting abilities. Baars et al. also suggest that there may be some form of the output editing process that occurs independently of the lexical/nonsense status of the context. The error rate for nonsense and lexical outcomes in a nonsense context as well as the error rate for lexical outcomes in a lexical context is more or less constant, but the error rate of nonsense outcomes in a lexical context drops significantly. These results seem to demonstrate how the editing mechanism targets and corrects the nonsense words into lexical words so that they fit the lexical context. Though this editing process may not correct all outputs to become lexically valid, the process may "correct" an error into a lexical item if it is in a lexical context and so expected to be lexically valid.

A possible semantic and phonotactic editing mechanism in the prearticulatory phase of speech encoding has also been proposed (Motley & Baars, 1976). The basis for such proposal comes from their evaluation of the semantic characteristics of targeted and spoonerized word pairs for semantic appropriateness. In their study, Motley & Baars (1976) observed that semantic and phonological interference of a targeted word pair together produce a greater amount of spoonerisms than phonological interference alone.

From this observation, they inferred that the semantic interference acts on the mechanisms in the prearticulatory decision-phase of speech production. The semantic editing mechanism, as they suggest, may act as a feedback or feedforward loop to check for semantic legitimacy and appropriateness of the selected word or phrase about to be uttered. A failure of this editing process could produce an utterance that does not semantically fit in the linguistic context, and overactivation of the mechanism could "correct" a selected word or phrase incorrectly so as to produce an error like a spoonerism.

Another problem these error editing mechanisms present to generative grammar is the involvement of time in these language processes. A grammar may be able to describe the steps taken to produce a spoonerism (or any language output), but it does not describe the steps taken in real-time. Error editing mechanisms show that encoding and decoding occurs at the same time—the mechanism must decode the initially-encoded linguistic information at each level (phonological, lexical, semantic, etc) to check for correctness and then recode it for the next step in production. While a grammar cannot show these steps' simultaneity because grammars are linear and categorical in nature, the biological human brain is capable of performing such simultaneous and gradient actions.

Instead of computationally based traditional models with a linear and categorical generative grammar, a language model based on neurocognitive architecture can account for this boundary blurriness as it could model how language moves from a physical process (of sound waves and electrical signals) to a cognitive perception along a continuum of information. Such a model would need to address and explain how the brain is able to interpret and superimpose segmentation onto unsegmented information with unclear boundaries. The brain somehow stores language processing into categories for efficiency, possibly the same or similar categories that we think of and perceive when processing language, all the while being able to quickly and easily synthesize and integrate complex information back into its continuative form.

Chapter 5: The Neurological Perspective

While more recent theoretical models have been moving away from definitive localization of function, the idea that certain areas of the brain seem to be associated with language processing has a lot of support through research in neuroscience. Recent advancements in neuroimaging allow for clearer, more accurate data collection. Technological advancements in data collection and distribution through the internet provide an abundance of resources and materials that can facilitate new experiments and encourage replication. Many academic institutions are beginning to establish and support brain and language labs with a variety of focuses from language acquisition and bilingualism to language-related neural development in children to the brain and sign language.

However, even before the developments in modern neuroimaging and other computerized research tools existed, neuroscientists performed experiments to study language in relation to the brain. One common way neuroscientists could study language processing was by studying language deficits associated with head-trauma or developmental disorders (UNC, 2016b). There are a number of documented language disorders that affect different components of language processing with different expressions (UNC, 2016a). One of the most well-documented types is aphasia, a condition generally defined as a neurological disorder that impairs the expression and comprehension of all language forms, resulting from damage to the portions of the brain responsible for language (NIH, 2015). Aphasia usually manifests quickly when the brain damage is caused by a stroke or a head injury, but progressive neurodegenerative disease or a tumor may cause a slow onset and progression of aphasia (NIH, 2015).

There are two main types of aphasia: fluent and non-fluent aphasia (NIH, 2015). Fluent aphasia is often caused by damage to the temporal lobe. Damage to Wernicke's area, a specific area along the superior temporal gyrus, causes the most common type of fluent aphasia called Wernicke's aphasia. Fluent aphasics are often able to produce long, syntactically accurate sentences, so they appear to be speaking fluently despite the fact these sentences typically have little comprehendible meaning. Fluent aphasics often appear unaware of their spoken mistakes, and they also have difficulty understanding the meaning of words or phrases produced by others. Like Wernickie's aphasia, damage to Broca's area along the inferior frontal gyrus leads to a type of nonfluent aphasia called Broca's aphasia. Characteristics of nonfluent aphasia generally consist of the ability to understand speech and to produce meaningful words (particularly content words like nouns and verbs), though production is often difficult, lacks function words, and forms ungrammatical sentences.

Some other common types of aphasias with specific behavioral effects include conduction aphasia, characterized by the ability to speak fluently but expressing difficulty in repeating words or sentences heard, and anomic aphasia, which is characterized by difficulty in naming objects while knowing what the object is (NIH, 2015). The relative localization of damage that appears to directly affect specific areas of language processing—along with the wider encompassing effects of more generalized brain damage as seen in global aphasia—is significant because it indicates not only that specific language functions may be generally localized to certain neural areas but also what functions might correlate with which areas. Knowing which areas of the brain are involved in certain language functions could better inform the language model. A more specific and accurate language model in turn could help improve the treatment and therapy given to patients with language pathologies by providing greater specificity about the treatment needed according to the pathology type.

General research on the functional neuroanatomy associated with language processing has also helped establish a basis for neurolinguistic research. As a whole, language processing can be broken down into two main components plus an intermediary stage. The first main component is linguistic information input, which involves the perception, recognition, and comprehension of linguistic input. The second main component is linguistic output, also known as language production. Language production may be expressed verbally through speech, graphically through writing, or visually through sign language. The intermediary stage of language processing is the storage of linguistic information that is first encountered as an input and then accessed and retrieved during linguistic output.

Language processing of linguistic information input, such as a single word, can be further broken down into two major stages: the first stage is the recovery of phonological information, and the second stage is the access to lexical and semantic information pertaining to that one word (Hickok, 2009). Certain areas of the temporal lobe have been implicated as being significantly involved in these two stages of language input processing. Both the right and left Superior Temporal lobes (STL) are thought to be involved in speech sound recognition; specifically, the Superior Temporal Sulcus (STS) has been identified as a critical site for phonological processing (the first major stage for speech input processing). The posterior temporal lobe areas (particularly the STS) seem to be more involved in phonological processing for auditory comprehension. The anterior area of the STS demonstrates particular activation in response to phonologic perceptual speech tasks, though this anterior portion of the STS is probably also involved in other aspects of speech perception like syntax or prosody processing.

Further studies of the STS suggest that it is an important site for representing and processing phonological information. Functional imaging studies contrasting speech stimuli with complex non-speech signal stimuli (to isolate phonological processes in perception) demonstrate activation along the STS (Hickok, 2009). This phonological processing of speech sounds appear to be left STS dominant, though lesions and imaging results suggest some sort of bilateral organization. However, bilateral organization does not necessitate symmetrical organization. The asymmetry of the phonological processing systems indicate that there may be parallel pathways involved in processing sound into meaning for spoken word recognition. The importance of these studies is that they suggest a functional boundary of language processing at the phonological level, anteriorly by the most anterolateral part of Heschl's gyrus (containing the primary auditory cortex) and posteriorly by the most posterior part of the Sylvian fissure.

The second major stage of language input processing is accessing lexical and semantic information (Hickok, 2009). While semantic processing is a major stage, there is disagreement among researchers as to the location of this processing. Some researchers believe conceptual information representation (like semantic information) is distributed throughout the cortex. They argue that when these representations activate, the same sensory, motor, and supramodal cortical systems are involved as when the representation was first processed (when the information was first learned). This means that when semantic information is first learned it is processed in a distributed manner across many areas of the cortex. Then the same systems distributed across the cortex are activated any time that information is accessed again for language input or output. On the other hand, some researchers believe semantic information is organized anatomically in a more localized area of processing in the anterior temporal region. Other researchers believe that semantic knowledge is organized even further into functionally specialized neural systems. Existing evidence implicates the posterior lateral and inferior temporal regions as important regions involved in converting sound information into meaning. The anterior temporal lobe may also be involved in semantic processing, but there may be evidence that the anterior temporal lobe is involved in more general activity rather than being specifically involved in linking sound input to meaning.

This disagreement among neuroscientists regarding the level of functional localization parallels the disagreement among linguists regarding whether or not contextual information like semantics is integrated throughout language processing or if it is separated and local to a processing module, particularly seen in the development of the cohort model (Marslen-Wilson, 1987). The disagreements among both neuroscientists and linguists further reflect the uncertainty of to what degree spoonerisms are isolated to a single domain or to several domains.

Yet, conceptual-semantic processing as a whole might not actually be unimodal—instead, it may actually involve supramodal representations of conceptual knowledge of objects (Hickok, 2009). Patients with semantic dementia have difficulty accessing object knowledge from both auditory language input *and* visual input, which would require the impairment of some cross-modal information integration process. One possibility is that the posterior lateral and inferior temporal lobe is involved in the acoustic processing of semantic knowledge while the anterior temporal lobe is involved in integrating the acoustic semantic knowledge with visual input. The challenge to understand language processing as a biological system capable of seamlessly integrating two types of information is also present in phonological-semantic information integration. For example, phonological information may be processed by systems in the superior temporal lobe while semantic information may be processed by systems in cortical regions outside of the superior temporal lobe. Spoonerisms demonstrate this type of integration as the semantic constraints on spoonerisms involve top-down processes that affect the bottom-up phonological construction of the spoonerism. The neurological support for this dual stream processing, particularly in which semantics is supramodel and so constitutes a top-down process involving context, provides support to the problems already presented to generative grammar: generative grammar cannot describe or explain how rules and constraints integrate to produce spoonerisms (or any language output) because rules and linguistic levels in grammars are traditionally modular, but the neurological evidence indicates that the brain is not modular. Thus, spoonerisms exceed the abilities of the grammar due to its nature as distributed knowledge and integrative processing.

In the end, understanding where semantic processing occurs—whether distributed or localized—is important because it can help us understand the level of localization or distribution of language processing throughout the cortex and to what degree language processing is integrated between each component (including the phonetic, semantic, and syntactic components). A system like this could help explain how two types of unimodal information are integrated to form a perceptual whole, a question neuroscientists have long investigated.

The second major component of language processing is language output production (Hickok, 2009). Like language input processing, language output processing for one word can also be divided into two major stages. The first major stage is the selection of a lemma (the appropriate lexical item intended) and the second major stage is the access of the lemma's phonological form and sound structure. This two-stage process in language output production means that there are also two areas in which output errors can occur: at the lexical level when selecting the proper lemma, and at the phonological level. Various types of speech production errors, like spoonerisms, suggests that there are these two major stages of language production, similar to the two major stages of linguistic information input.

How a lemma is selected in language output is often a concern in language production models, though it is generally approached through the focus of phonology and how phonological information is assembled to construct the appropriate lemma. The posterior language cortex in the left hemisphere appears to be significantly involved in speech production on the phonological level (Hickok, 2009). In fact, auditory input seems to have an important influence on speech production output that (as adult-onset deafness indicates) helps maintain articulatory tuning on phonetic, pitch, and phonemic sequence production processes. This input influences first encodes the language input stimulus (like a spoken word or phrase) into the phonological auditory system before it is mapped onto the corresponding motor articulatory sequence, through which the sensory-representation of the stimulus word form is learned. The motor articulatory sequence is then consolidated as a learned motor unit that requires little sensory guidance in future activation. Damage to the dorsal posterior superior temporal gyrus (STG), to the supramarginal gyrus, or to both the STG and the supramarginal gyrus produces speech production deficits, particularly conduction aphasia in which auditory comprehension (input processing) is relatively good but speech production (output) is poor. These speech production deficits are likely from a deficit in the sensory-motor integration system for speech.

In terms of the auditory sensory component of this sensory-motor integration system, the Sylvian fissure at the parietal-temporal boundary (Spt) has been implicated (Hickok, 2009). The Spt is an area in the left posterior planum temporal region that appears to be distinct from the spatial hearing-related functions of the other, more anterior portions of the planum temporale. Research has shown its integrative function of sensory-motor phonological information, indicating the speech sensory-motor integration system is likely to be impacted by damage to this area as well. In particular, the left posterior superior temporal regions are implicated in general speech production. The posterior part of the left planum temporal region activates during picture naming tasks and demonstrates length effects, frequency effects, and a time-course activation that are consistent with the naming task's phonological encoding stage. Lesions in this particular area are associated with conduction aphasia. Likewise, if these areas—the Spt, STG and/or the supramarginal gyrus—also show abnormal activation during spoonerism production, it would evince that there is a sensory-motor speech integration process in this area of the brain. Not only that, but such activation would suggest that spoonerism production may be a benign form of sensory-motor integration error and so may provide a new perspective through which to study aphasia.

To address the problem of how sensory and motor information interact during speech production, the classical model consulted is Geschwind's model indicates the arcuate fasciculus (a white matter association pathway) as a connection between semantic information in Wernicke's area and motor and morphemic information in Broca's area (Hickok, 2009). A more recent model for sensory-motor integration is a cortical integration network for speech and speech-related abilities with properties like sensory-motor integration regions in the visual system's dorsal stream, connectivity with frontal motor systems (specifically motor-effectors), and multisensory responses. This cortical integration network includes the Spt, which has been argued by some researchers to support sensory-motor integration for speech/vocal tract effectors because of the similarities in the response properties to IPS (intraparietal sulcus) areas. Support for this comes from fMRI studies that show activity in the Spt during both perception and production of speech, which seems to suggest that the Spt is functionally connected to motor speech areas and is organized around the vocal tract effector system. However, other studies do indicate that the Spt may not be speech specific, as it is also sensitive to speech-related visual stimuli like silent lip-reading, and to non-speech auditory stimuli like melodic humming. Regardless of if the Spt is speech specific or not, damage to the Spt produces sensory-motor deficits but not speech *recognition* deficits, so it is possible that the Spt is involved exclusively in speech production output and is not involved in speech input recognition.

Based on the anatomical and functional organization for language input and output processing seen above, researchers propose a dual stream model for phonological processing of auditory information along a similar path as the dual visual stream, where the asymmetric bilaterally organized ventral stream is involved in speech comprehension while the left-dominant dorsal stream (involving the Spt and the posterior frontal lobe) is involved in converting speech signals into articulatory representations in the frontal lobe (Hickok, 2009). This neurological model harks back to cohort and TRACE language production models. While Marslen-Wilson (1987) in cohort argues for a non-interacting dual stream (an aspect of the dual stream the neurological evidence does not support), Goldrick et al. (2016) demonstrates how the dual streams do interact and affect each other behaviorally. In TRACE, McClelland & Elman (1986) favor an interacting dual stream. These linguistic models are based on computational processing and so cannot completely model natural biological language processing; however, this neurobiological dual steam model shows how the general framework these earlier models use can be restructured to begin the formation of a language model that accounts for the grammatical, behavioral, and neurobiological components behind at least spoonerism production, if not all language production.

Chapter 6: A Primary EEG Study

Through discovery of the apparent rules and conditions for spoonerisms to occur, a few techniques were developed to induce spoonerisms as naturally as possible in a controlled laboratory setting. Attempting to manipulate natural spontaneous speech to test theories on almost any sort of speech and language processes comes with multiple problems and challenges that need to be addressed. The first major challenge is controlling and manipulating speech production without participants knowing that their speech is being guided in some way (or at least knowing the targeted outcome). If participants know what outcomes researchers are targeting, then the speech outputs will probably be influenced by the knowledge in some way, consciously (such as trying to produce or avoid targeted outcomes) or unconsciously. Secondly, it is difficult to control all the independent variables that can influence a speaker's output, so it is difficult to isolate and identify which individual factors directly cause, or at least in some way influence, a specific aspect of an output and which factors do not. Overall, the biggest challenge that researchers have had to consider in attempting to elicit spoonerisms in a laboratory setting is how to isolate and control the one independent variable they want to manipulate without affecting any other influencing factors.

Experimentation on spoonerisms began not with a spoonerism-specific study, but with a study on motor stress pre-entry with prosodic stress. In Donald MacKay's experiment, participants were asked to rapidly repeat sequences of four syllables "tay," "gay," "bay," and "day," with prosodic stress on a certain syllable designated by the experimenter (1971). The study found that participants would accidentally switch the stressed syllable with the preceding syllable. While not explicitly a study on spoonerisms, MacKay's study provided a start for Motley and Baars' design for eliciting phonemic switches between adjacent words called the "Spoonerisms of Laboratory Induced Predisposition" (SLIP) method (1976). In this procedure, experimenters first compiled potential spoonerizable utterances as targeted word pairs, such as "trail head" and "billing flow." Then they primed participants to spoonerize the targeted word pair by preceding each pair with another word pair that is resembles the word pair whose phonemes are switched into the orientation of the target spoonerism. For example, "**t**rail **h**ead" could be preceded by "**h**ail **t**rend" and "**h**am **t**end"; "**b**illing **f**low" could be preceded by "**f**oe **b**linking" and "**f**linging **b**ow." The participants were asked to read the word pairs that were presented to them one at a time, but they were asked to verbalize only the word pairs that were cued auditorially with a buzzer (the word pairs targeted for spoonerizing). The reason participants were to verbalize only the targeted word pairs was to limit the likelihood that any verbal mistake elicited was from articulatory confusion rather than other potential non-motor causes that are being tested. To make sure that the priming word pairs actually primed the participants rather than participants paying attention only to word pairs with auditory cues and ignoring word pairs without an auditory cue, experimenters told participants that they needed to remember all the word pairs for a memory-recall test administered later. Then, to prevent participants from predicting correlation between sound cues and targets as well as to prevent anticipation of the sound cue, several "neutral" word pairs (non-targeted and non-priming word pairs) were presented, some of which received randomly distributed auditory cues.

Even when accounting for some of the challenges inherent to trying to elicit a certain phenomenon while retaining the participants' natural spontaneity of speech, there were some issues with the SLIP method. First, it is uncertain if the SLIP method elicits slips in speech output (decoding) or slips in information input (encoding), an important distinction for understanding where in the speech process spoonerisms occur. Second, the method does not reveal the role of articulatory interference in the spoonerism elicitations. If articulatory interference and confusion are what cause spoonerisms in the SLIP method, then these laboratory-generated errors are more of a motor-error (like tongue twisters) than naturally occurring spoonerisms. Finally, the use of primers to elicit spoonerisms may not accurately reflect the real cause of natural spoonerisms since natural spoonerisms are spontaneous and most are not primed with interference from preceding word pairs. Overall, the unavoidable conditions of a laboratory-based experiment can affect the process of natural speech production, thereby inadvertently influencing the results.

A few researchers skeptical of the SLIP method suggested that the methodology as a whole should be reevaluated. Sinsabaugh and Fox (1976) call for a critical reevaluation of Motley and Baar's 1976 data as being open to many different alternate interpretations when they performed their replication of the SLIP method. Sinsabaugh and Fox stated that their replication produced non-spoonerism speech errors more frequently than spoonerisms and suggested that many of the other speech errors they found resulted from memory confusion instead of elicitation from actual spoonerisms. The types of speech errors that were more frequent in their replication were a failing to verbalize a response, an errored response that was phonetically unrelated to the preceding word pair (and so the error seemed to have been uninfluenced), a response that was phonetically unrelated to the targeted word pair, and a response that included one or both of the priming word pairs instead of the targeted word pair or its variations. As they attempted to identify flaws in Motley and Baas' procedure, they claimed that other speech errors were far more common because spoonerisms only made up a small fraction of the total speech errors, which they explained to be caused by proactive inhibition or acoustic confusion in short-term memory.

While Sinsabaugh and Fox provide important counter-arguments against the SLIP method and questions about the factors and conditions that are thought to elicit spoonerisms in the SLIP method, Motley (1986) responded to the critiques of Sinsabaugh and Fox with a re-replication of the experiment using the SLIP technique. In his rebuttal to Sinsabaugh and Fox's criticism of the original paper he coauthored with Baars, Motley analyzed the data and interpretations Sinsabaugh and Fox published before providing his own re-replication data. Motley's re-replication supports the original hypothesis that the frequencies of lexically legitimate spoonerisms will be significantly greater than the frequency of lexically illegitimate spoonerisms using the SLIP technique (Baars & Motley, 1975). Motley (1986) stated that Sinsabaugh and Fox's results could be caused by improperly executing the procedure in ways that would produce many other verbal slips aside from spoonerisms. For example, placing the priming word pairs the farthest from the target word pairs instead of closest to the target can weakening the bias for the spoonerized form, thereby reducing spoonerism production (Motley, 1986). Other possible variables that could have reduce spoonerism production include the use of a computer screen to present word pairs rather than using a memory drum, the possibility participants were aware that the task was to elicit spoonerisms or were able to predict cues, or the possibility that Sinsabaugh and Fox presented the cue for the targeted word pair simultaneously or too soon after the stimulus (Motley, 1986).

In support of the SLIP method, Motley (1986) also explained that not only had he and his colleagues reported over twenty experiments using the SLIP method to test speech production, but other researchers have successfully used the SLIP method in experiments. Moreover, the SLIP method has continued to be used by a variety of researchers. Though the number of different experiments conducted by different researchers using the SLIP method does not necessarily mean that the SLIP method is without flaws or should not be critically reevaluated, the numbers do suggest that the SLIP method does have at least some sound methodology that can produce reasonable and reliable results.

Despite the SLIP method's apparent success at eliciting spoonerisms in a controlled setting while retaining a sense of natural speech production, there have been relatively few recent studies using the SLIP method, and SLIP method-based research using modern neuroimaging is especially lacking. For example, EEG (electroencephalographic) studies using the SLIP method could provide a base of information useful for understanding spoonerisms as a neurobiological occurrence. The purpose of conducting such a study in this paper is to test if there is a strong correlation between verbal speech error behavior and the results of one previous SLIP method-EEG study (Möller et al., 2007) by first eliciting spoonerisms and then measuring how long post-stimulus a correlating ERP (event-related potential) occurs. These results could then help indicate where in the language encoding-to-production process the error that produces this kind of verbal slip occurs. Spoonerisms have not been widely studied from a quantitative neurolinguistics standpoint as most research has been conducted from a behaviorally-descriptive psycholinguistic perspective, so this experiment would help contribute quantitative data to the field and support previous studies that focus on the neuroanatomical aspect. The overarching purpose of this research is to better identify the neural mechanisms that produce speech errors in order to better understand language deficits with neurological causes as well as improve understanding of normal language processes.

If there is a correlation between relatively specific localized neural areas and spoonerisms, this connection could provide insight into how language (phonological, morphological and/or lexical) encoding and retrieval normally occurs by implicating the mechanisms involved in language errors production. The findings in a study by Möller et al. (2007) seem to indicate that at least some articulated sound errors are preceded by competing representations of articulation in correlation with SMA (supplementary motor area) activation, suggesting that spoonerisms reflect an interruption of speech production in the early stage and so are not exclusively a semantic or even phonological phenomenon. Findings in phonological processing errors demonstrate that ERPs correlate with the superior temporal sulcus region, whereas findings in semantic studies demonstrate that ERPs occur later and correlate with a wider area throughout the cortex

(Hickok, 2009). Because the Sylvian parietal-temporal (Spt) regions have been implicated in general speech production and perception (Hickok, 2009), and in the context of earlier ERP results (Möller et al., 2007), it seems likely that spoonerisms are not exclusively a phonological, semantic, or syntactic error but more a symptom of a glitch in an even earlier stage of speech production, a glitch that impacts and informs these different components of language. This experiment is to determine if the findings of Möller et al. (2007) can be supported and determine how the findings of these left anterior negativities affect the current language-production model.

I hypothesize that spoonerisms will produce an increased negative ERP response that correlates in time to an earlier, more integrative neurological process rather than those correlated with exclusively phonological, semantic, or syntactic deviation. There are a few questions I hope to answer through this experiment: where between encoding and decoding processes do spoonerisms occur; are spoonerisms an error at a phonological, morphological, or lexical level; and are spoonerisms a conflict between a top-down and bottom-up process where the cognitive idea of what to say conflicts or glitched the processes that actualize the idea into a linguistic output?

Method

The study included 20 participants who were 18-50 year-old native English speakers without any known speech pathologies, uncorrected vision, or uncorrected hearing. Anyone outside of the age range, who was not a native English speaker, who had a speech pathology, uncorrected vision, or uncorrected hearing was excluded from the study.

Participants were fitted with a 16-electrode headset to measure and record ERPs (Event-Related Potentials) through a Cyton Biosensing board[3]. These dry electrodes were placed directly on the scalp at the locations Fp1, Fp2, F7, F3, F4, F8, T7, C3, C4, T8, P7, P3, P4, P8, O1, and O2 on the International 10-20 system (Appendix A). Participants were presented with word pairs on a computer screen (Appendix B), on which they are asked to keep their fixation central. Their task was to vocalize a target word pair as fast as they could immediately upon hearing the response cue presented after the onset of the target pair. There were 2-7 word pairs per trial: one pair was the targeted word pair and the rest were word pairs intended to prime spoonerism production. At the end of each trial was a memory task in which a single word from the preceding series of word pairs was presented for 6 seconds and the participant was asked to recall the corresponding missing word by saying the completed word pair out loud. The memory task was used to both ensure the participants were reading all word pairs fully and to disguise the targeted word pairs. There were three sets of 25 trials (75 trials total), with a total length of time of 60 minutes. Participants' verbal responses are recorded with an audio recorder to classify each response type (full spoonerism, partial spoonerism, other verbal error, no error), and ERPs were recorded from the scalp and recorded offline.

---

[3] The author would like to thank the Regis CC&IS and the Data Sciences Department for access and use of the OpenBCI EEG equipment

Results

The ERP results of this experiment are inconclusive. Altogether, participants produced 9 full spoonerisms (Table 2; with three participants spoonerizing twice and three spoonerizing only once) and 14 partial spoonerisms (Table 3; three participants spoonerized once, two participants spoonerized three times, and one participant spoonerized five times). Responses were considered partial spoonerisms when one out of the two words exhibited the targeted phoneme even if participants corrected their speech mid-response. Self-corrections are indicated by dashes.

Table 2. ***Fully spoonerized verbal responses of the targeted word pair and the corresponding priming word pairs***

Targeted Word Pair	**Snoring Boar**	**Cook Goes**	**Lame Fate**	**Rig Bisque**	**Right Mead**	**Bind Wink**	**Chart Hunk**	**Yarn Bard**
Spoonerized Response	Boaring Snore	Goo Croes	Late Fame	Big Risk	Mighty Read; Might Read	Bink Wind	Chunk Heart	Yard Barn
Priming Words	Billowing Sheep Buy Garb Uninteresting Sleep Boarding Snow	Deep Keys Good Gore Goopy Clothes	Super Star Tardy Time Fan Sun Fail Late	Big Risk	Safe Road Mean Rise Maybe Read	Warm Blanket Wind Blink Wash Bin	Happy Children Heart Chunk Home Choice	Tree Bark Yarn Bard Lip Balm

Table 3. ***Partially spoonerized verbal responses of the targeted word pair and the corresponding priming word pairs***

Targeted Word Pairs	**High Top**	**Billing Flow**	**Bird Feeder**	**Dart Board**	**Fruit Fly**	**Mad Bug**	**Rig Bisque**	**Found Rind**	**Bind Wink**	**Map Nook**	**Barn Yard**
Partial Spoonerism	Hip Hop	Fl-- Billing Flow	Fird Feeder	Bart-- Dart Board	Flute-- Fruity Fly	Bad Bug	Risk Bisque	Round Rind	Bink Wink	Nap-- Map Nook	Yar- Barn Yard
				Bart-- Dart Board		But- Muddy Bug	Big Bisque				
Priming Word Pairs	Tip Top	Great Abundance	Flouncing Blue	Bad Goof	Flag Fraud	Coffee Cup	Big Risk	Free Ring	Warm Blanket	Marker Case	Tree Bark
	Tie Hop	Filling Bow	Faux Fur	Busy Duck	Flat Freight	Big Date		Flounce Behind	Wind Blink	Short Sleep	Yarn Bard
		Flinging Blow	Roof Tops			Runny Mud		Round Find	Wash Bin	New Moon	Lip Balm
			Food Blender							Noodle Mush	

Based on time calculations, a few potential ERPs that may correspond with the spoonerized responses were identified (Figure 2). However, statistical analyses of the EEG data recorded could not be performed for a number of reasons, mainly due to the lack of time-locking between the data stream, the stimuli presented, and the verbal responses given.

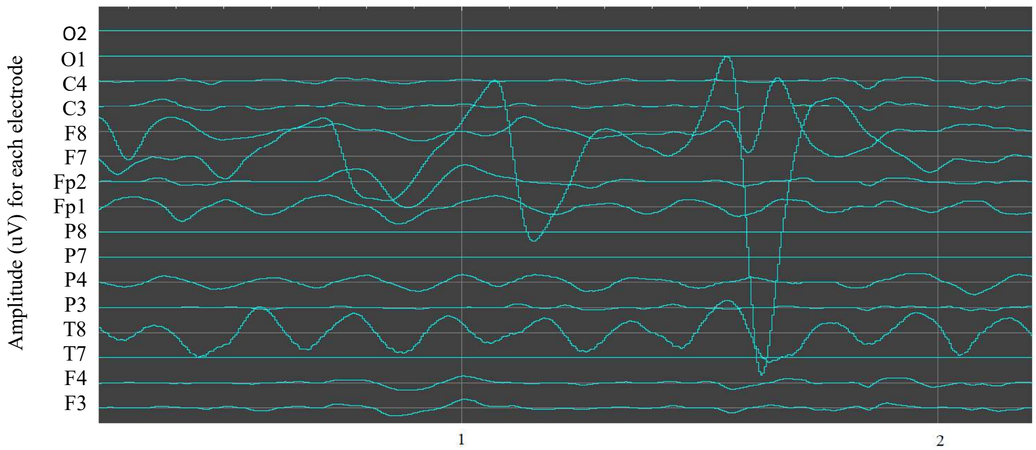


Figure 2. **EEG montage showing amplitude (uV) measured by each electrode over a 2 second period potentially in response to a sound cue and spoonerism production. This montage demonstrates a potential ERP detected and recorded by the F7 electrode in response to a sound cue and then to the production of the spoonerism "Boaring Snore" (the targeted word pair of which was "Snoring Boar"), starting 0.2 seconds pre-sound.**

Discussion

A number of factors contributed to inconclusiveness. To start, there were a few problems with the hardware used, mainly involving the electrodes' inability to record signals. For some participants, not all the electrodes on the headset could touch the scalp because of the size and shape of the inflexible headset (which was 3D printed with hard plastic). Some of the electrodes were unable to receive a signal for other reasons, such as possible misconnections. The physical electrodes did not correlate with the channel names on the OpenBCI GUI (which means that when the headset was assembled, the electrodes were misarranged), so a key was required to discern which electrode was receiving a specific signal and where on the array the signal was coming from. There were also a number of software problems. Sometimes only part of the data stream was recorded or was transferrable to the EDFbrowsing program used to view and process the data. There was also a lack of time lock between the stimuli presented, the behavioral responses, and the ERPs recorded. As such, the ERPs identified as correlating with the spoonerisms produced are based on time calculations, and so there is uncertainty if the ERPs identified do in fact represent the spoonerisms produced.

Time limitations also contribute to the inconclusivity of the results. Because of time and resource constraints, there were not nearly enough trials with each participant (there were only 75 instead of the originally intended 550 sets of trials) and there were not nearly enough participants for enough spoonerisms to be produced, let alone analyzed for significant ERP results. Ultimately, my own learning curve reduced the efficiency of the experiment as I had to learn how to design such an experiment, how to use the EEG headset and the OpenBCI GUI system, how to use the Matlab script converter in Octave to convert the files the OpenBCI recorded the data into an ASCII file, how to then convert the ACII file to an EDF file, and then finally how to use the EDFbrowser software to view and analyze the ERP data.

In addition to the factors that led to inconclusive data, improvements could also be made to the methodology in respect to modifying the SLIP method for modern technologies. The timing of each trial may not have been optimal for spoonerism production: the amount of time each word pair was presented to the participants may have been too long or the amount of time in between each word pair presented may have been too long and so may have affected the priming effect or the likelihood of a spoonerized response; on the other hand, the amount of time between trials or between sets of trials may not have been long enough as many participants reported feeling overwhelmed by the rate and amount of words they focused on. Perhaps having more sets with fewer trials (such as 7 sets of 5 trials) along with more time between sets would help alleviate stress while retaining the speed of response necessary for mimicking natural, spontaneous speech.

Certain aspects of the memory quizzes may have also complicated the procedure. While some interesting patterns arose from the memory quizzes (such as completing the word pair so that the two words rhymed with each other or that the completed word pair rhymed with the previously verbalized word pair), these patterns were not the targeted focus for the study. Moreover, participants seemed more focused on recalling the missing word for the memory quiz after each trial than saying the cued word pairs out loud. While this may be an advantage for the methodology because it seems like it would make participants less guarded in their speech, fewer spoonerisms may have been produced because participants were particularly focused and careful in their responses because they felt like they were in a test setting.

Though inconclusive, this experiment still holds value. In many respects, it was a pilot test using the SLIP method with modifications for modern technology and a mostly new list of target and priming words. What is more, the process of finding, adapting, and applying research methods taught me a lot as an undergraduate researcher. It has taught me the amount of time and work required to set up and run an experiment; I became familiarized with using an EEG headset and the associated software; it provided me with experience working with human participants; it exposed me to ERP data processing.

Despite the inconclusive nature of these particular results, future research should be conducted. SLIP methodology with better adaptations that make use of modern technology in conjunction with more precise and accurate equipment—such as with an EEG headset with a full array of electrodes or an fMRI—and a time-locked system could yield more conclusive data.

More specific areas to explore within the neurobiology of spoonerisms could include testing to see if spoonerism production is correlated to working memory (such as how rhyme seems to prompt participants to respond in a certain way to the memory quiz questions) or to test to see if spoonerisms are involved in a process that occurs even before production begins (such as in an encoding process). Anatomical studies could analyze structural differences between individuals to see if there is any correlation between frequency of spoonerism production and neuronal, glial, or dendritic densities in certain regions of the brain. A longitudinal study could survey the frequency of spoonerisms over the various stages of cognitive development and decline, the results of which might not only further understanding of spoonerism production but also language acquisition throughout development. Though this experiment may not directly contribute to the fields of linguistics or neuroscience, it perhaps exposes the severe lack of research existing in this important intersecting area. Neuroimaging that records neuronal activity during spoonerism production can still tell us a lot about language processing by showing where this early-stage error occurs both neuroanatomically and in the language production process.

Chapter 7: Larger Application and Conclusion

Why Are Spoonerisms Important?

Many might believe spoonerisms are just a random oddity, an insignificant slip of the tongue in a careless speaker. But spoonerisms are significant, in part because of their oddity. When linguists describe the process of metathesis, it is not often under the perception as a form of error, but rather just as a normal language process. Metathesis as a normal language process is generally studied in the context of diachronic historical linguistics. For example, the /l/ and /r/ from the Latin word "pa**r**abo**l**a" to "pa**l**ab**r**a," or the vowel sounds and the /r/ in Old English's "beorht" in Old English metathesized to "bryht" in Middle English ("bright" in modern English). If spoonerisms are a form of metathesis and are a form of verbal error, are all metatheses technically errors? Or do spoonerisms just bridge the gap between language "errors" and language "changes"? Some may argue the latter, that language changes are not language errors and errors are not changes. But perhaps errors and changes are not really so different, perhaps language "errors" and language "changes" are both types of language variation. So why does this distinction matter? Why does it matter if we consider "errors" and "changes" and "variations" to be different events or the same occurrence? These seemingly subjective questions reveal the dynamic, ambiguous nature of language and our struggle to categorically confine it within set, distinct boundaries.

By bridging the gap between language error and language change, spoonerisms also bridge the gap between normative and pathological language processing. They are normative in the sense that they are a common occurrence produced by individuals conventionally considered neurologically healthy; they are pathological in the sense that they are unintended disruptions of intended speech production, and sometimes they interfere with a speaker's production process so much that the speaker has significant difficulty trying to correct and produce the proper, originally intended utterance. This nature of spoonerisms as a form of benign speech error provides a unique and rich window into the complexity of language processing. They arise in our everyday communication spontaneously and unintentionally. We can use them intentionally for humor or to draw attention to particular ideas and concepts behind a certain wording. They offer a focal point for language processing models as they challenge models to adequately describe a range of language production behaviors (from "normal" to "abnormal") and support analysis of existing language processing models and theories. Spoonerisms are fairly simple to incorporate in a variety of experimental designs, including correlations with neural systems, thereby bridging the gap between language study in linguistics and neuroscience through a measure of empiricism.

For me, this study of spoonerisms has been an important start to further study of how language and the brain interact. I would like to continue conducting research in a lab setting as well as in the field, working with different language communities, working with records of languages, or even just with natural and spontaneous conversation, in order to better understand language and establish a more empirically-based neural model of phonological acquisition in bilingual speakers. Practical applications for neurolinguistics in respect to language and the brain include clinical work with aphasics, for working for social awareness and acceptance of linguistic variations like bilingualism, or for cultural support work with communities facing language death which could then lead to loss of culture and identity. Neurolinguistics provides a way to address the personal component of neuroscience that often seems overlooked—the disorders discussed and tested (linguistic or otherwise) affect real people, and experiments concerning these disorders not only contribute to our understanding but can also tangibly help people affected by disorders. Ultimately, I want to not simply learn and contribute research, but to find a way to use such information to help people who face linguistic and psychological challenges.

References

Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning & Verbal Behavior*, *14*(4), 382-391.

Chomsky, N. (1957). *Syntactic Structures*. Berlin: Mouton de Gruyter.

Noam Chomsky in *Syntactic Structures* (1957)

Erard, M. (2007). Read my slips: Speech errors show how language is processed. *Science*, *317*(5845), 1674-1676.

Finely, S. (2017). *Perceptually based constraints and metathesis: evidence from artificial grammar*. Retrieved from https://journals.linguisticsociety.org/proceedings/index.php/amphonology/article/view/3985.

Fodor, J. A. (1983). Four accounts of mental structure. *In the modularity of mind* (pp.1 38). Cambridge, MA: MIT Press.

Frederickson, N., & Wilson, J. (1996). Phonological Awareness Training: a New Approach to Phonics Teaching. *Dyslexia (10769242)*, *2*(2), 101-120.

Fromkin, V. A. (1973). Slips of the tongue. *Scientific American*, *229*(6), 110-117.

Goldrick, M., Keshet, J., Gustafson, E., Heller, J., & Needle, J. (2016). Automatic analysis of slips of the tongue: Insights into the cognitive architecture of speech production. *Cognition*, *149,* 31-39.

Hickok, G. (2009). The Functional Neuroanatomy of Language. *Physics of Life Reviews*, *6*(3),121–143.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, *8*(3), 323-350.

MacKay, D. G. (1971). Stress pre-entry in motor systems. *The American Journal of Psychology*, *84*(1), 35-51.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71-102.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.

Möller, J., Jansma, B. M., Rodriguez-Fornells, A., & Münte, T. F. (2007). What the brain does before the tongue slips. *Cerebral Cortex*, *17*(5), 1173-1178.

Motley, M. T. (1973). An analysis of spoonerisms as psycholinguistic phenomena. *Speech Monographs*, *40*(1), 66.

Motley, M. T., & Baars, B. J. (1976). Laboratory induction of verbal slips: A new method for psycholinguistic research. *Communication Quarterly*, *24*(2), 28-34.

Motley, M. T., & Baars, B. J. (1976). Semantic bias effects on the outcomes of verbal slips. *Cognition*, *4*(2), 177-187.

Motley, M. T. (1986). On replicating the SLIP technique: A reply to Sinsabaugh and Fox. *Communication Monographs*, *53*(4), 342-351.

National Institute of Health (2015). *National Institute on Deafness and Other Communication Disorders: Aphasia* (Pub. No. 97-4257). Bethesda, MD: NIDCD. Retrieved January 19, 2018, from https://www.nidcd.nih.gov/health/aphasia.

Sinsabaugh, B. A., & Fox, R. A. (1986). Reevaluating the SLIP paradigm: A research note. *Communication Monographs*, *53*(4), 335-341.

University of North Carolina Health Sciences Library (2016a). *Adult Neurogenic Communication Disorders*. Retrieved from https://hsl.lib.unc.edu/speechandhearing/neurogenicsaphasia.

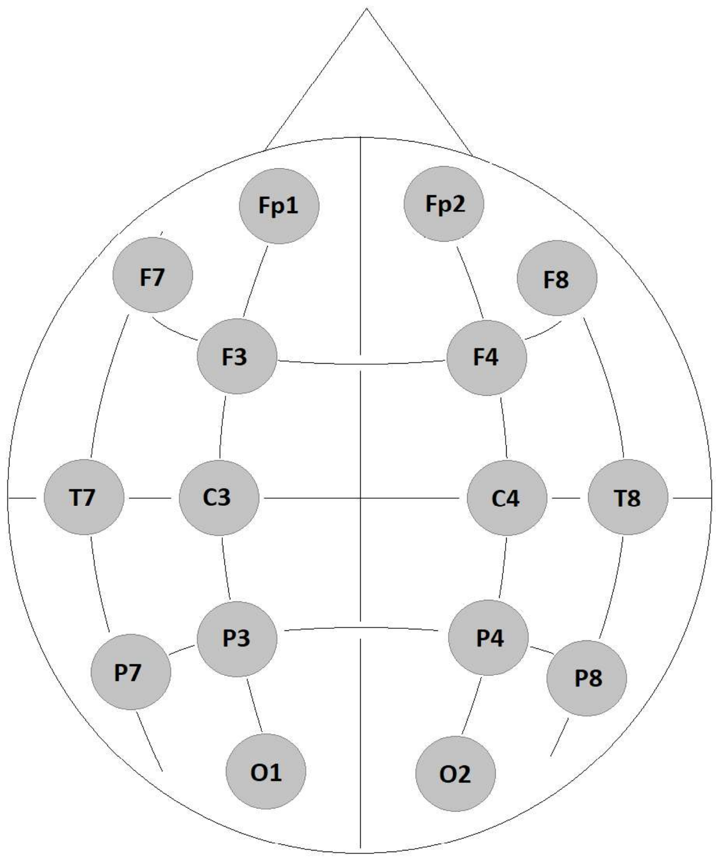
University of North Carolina Health Sciences Library (2016b). *History of the Professions: A Brief History of Audiology*. Retrieved from https://hsl.lib.unc.edu/speechandhearing/professionshistory.

Varvara, P., Varuzza, C., Sorrentino, A. P., Vicari, S., & Menghini, D. (2014). Executive functions in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*, 120.

Wickelgran, W. A. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review*, *76*(1), 1-15.

Appendix A

Location of the electrode placement along the International 10-20 system



Appendix B

Spoonerism Word Pair List

Rats Mice

Bus Stop

Stamp Blot

Trail Head

Hail Trend

Ham Tend

Grave Digger

Dane Giver

Nobel Laureate

Bat Tap

Pat Tam

Cat Nip

Door Raid

Car Toll

Luck Snack

Wall Hole

Jill Dean

Shoulder Sash

Buy Garb

Given Bin

Limb Dark

Dim Lark

Dune Buggy

Big Hook

Hard Line

Bard High

Learn Him

Newt Mine

Mute Nine

High Top

Tie Hop

Tip Top

Hot Mug

Mop Hog

Snoring boar

Boarding snow

Billing Flow

Filling Bow

Flinging Blow

Blinking Foe

Pleading Seed

Seeing Plenty

Seeking Pillows

Window Sills

Bird Feeder

Faux Fur

Flouncing Blue

Food Blender

Tour School

Skiing Too

Skewer Tool

Blue House

Cool Blouse

Hue Bloom

Big Risk

Rig Bisque

Make Clear

Clean Muck

Click More

Cake Mirror

Round Find

Found Rind

Flounce Behind

Free Ring

Blind Wink

Wash Bin

Warm Blanket
Wind Blink
Keep Blear
Beep Clear
Big Clunk
Black Chunk
Read Stew
Seed Rue
New Moon
Map Nook
Marker Case
Mushy Noodles
High Chair
Fluorescent Hair
Carnival Fair
Chart Hunk
Home Choice
Happy Children
Heart Chunk
Plowing Lakes
Purple Leaves
Looking Pleased
Last Bone
Based Low
Bear Lane

Blast Zone
Wash Pot
Plot What
Pilot Wing
Posh Parlor
Blue Words
Winter Boots
Warm Bread
Wooed Birds
Many Days
Delightful Maze
Daring Mays
Heavy Dog
Plush Couch
Clutch Purse
Cozy Jacket
Cushion Plant
Bare Cold
Happy Party
Perfect Hearing
Partly Cloudy
Pattering Heart
Fully Hearty
Tap Show
Big Toe

Sharp Talk
Shopping Cart
Sneak Loose
Leaky Shoes
Peeking Snooze
Fresh Salt
Session Taker
Seeing Farms
Barn Yard
Yearn Sing
Lip Balm
Tree Bark
Yarn Bard
Top Key
Copper Tank
Crystal Stone
Long Talk
Young Teen
Script Team
Tall Beam
Happy Feet
Fancy Gift
Flying Here
Flinging Heap
Dump Truck

Dizzy Block

Dark Buck

Bland Food

(From Baars, Motley, & Mackay, 1975)

Could Gore

Cook Goes

Deep Cot

Deed Cop

Keen Lap

Keys Lab

Dumb Seal

Dump Seat

Big Dues

Bit Dukes

Luke Risk

Bought Cat

Can sat

Call Bit

Lame fate

Fail late

Bad Goof

Dart Board

Busy Duck

Safe Road

Right Mead

Rise Mean

Fail Sun

Fate Sum

Lean Cap

Lead Cat

Met pile

Mess Pipe

Rail Seep

Raid Seas

Soul Rock

Soak Rot

Might Toss

Mice Taught

Bail Toss

Bait Tot

Taught Far

Long Rice

Log Ripe

Some Toys

(From Motley & Baars, 1976)

Pick Soap

Sick Pope

Ill Bishop

Stricken Priest

Tame Soon

Same Tune

Known Song

Similar Melody

Mice Knob

Nice Mob

Good Group

Pleasant People

Sat Feet

Fat seat

Large Chair

Chop Sticks

Fruit Fly

Flute Fry

Flat Freight

Flag Fraud

Light Rake

Right Lake

Pine Fig

Fine Pig

Fire Pit

Five Pills

Bad Sum

Meek Lad

Leek Mad

Mean Cut

Keen Mutt

Bad Mug

Mad Bug

Tall Boy

Big Date

Wage Rate

Rage Weigh