The Influence of Cognitive Factors on Category-Specific Phonology

Jonathan Manker

University of California, Berkeley

A growing body of literature has documented an array of phonological patterns---either static inventories or active rules---which are sensitive to the grammatical category of the words they govern. Here I consider grammatical categories as including both syntactic classes such as nouns, verbs, and adjectives, and broad classes such as functional and content morphemes. A parallel body of literature has shown psycholinguistic and neurological differences in how different grammatical categories are treated. These independent observations in both the phonological and cognitive literature suggest a currently unexplored question: Are the observed category-specific phonological asymmetries the reflection of asymmetries in the mental organization or processing of these categories? A review of the literature reveals little solid evidence that there are either innate biases in category-specific phonology or that categories form neighborhoods that facilitate the spread of phonological patterns in the lexicon, when other possible factors, such as syntax and semantics, are considered. However, observations from the dual-stream model of speech perception suggest there may be emergent differences in how more predictable, functional categories are processed as opposed to less predictable, content categories.

1. Introduction

In recent decades, linguists have begun to explore how many aspects of grammar interact with phonological patterns in language. In particular, studies coming from a wide range of approaches have documented and analyzed the effect of grammatical categories on sound patterns. Here I will use the term grammatical categories liberally to include classes based on syntactic function (nouns, verbs, adpositions, determiners, etc.), and the broader classes including content (nouns, verbs, adjectives) and functional morphemes (determiners, prepositions, inflectional affixes, etc.). While some studies focus only on open vs. closed class words, content vs. function
words, or root vs. affix distinction, I am considering all of these together, as excluding any of these classes from the discussion could weaken some of the generalizations that can be made.

From the literature on sound change, a number of studies (Crowley 1997, Donohue 2005, Blevins & Lynch 2009) have analyzed apparent cases of class conditioned sound changes (for example, X becomes Y when in the environment of word class Z) and have discussed whether it is necessary to invoke such a condition. Other studies investigating seemingly irregular, diffusive type sound changes (Phillips 2006, Brown 2013, Bybee 2014) have considered the role of grammatical categories in the gradual spread of sound changes as well as alternative, usage-based explanations. Lastly, works concerned primarily with synchronic phonology have appealed to grammatical categories in explaining surface asymmetries between sound patterns in roots vs. affixes (McCarthy & Prince 1995, Steriade 1995, Beckman 1998) as well as nouns vs. verbs (Smith 1997, 2010, McCarthy 2002, Cable 2003). These studies examine both the differences in inventories and allowable contrasts with respect to different categories, as well as class-based asymmetries in the application of phonological rules.

Meanwhile, other branches of research have developed in the fields of psycholinguistics and neurolinguistics. This research focuses on how different grammatical categories are stored and processed in the brain. Many studies have shown distinct patterns in the ways nouns and verbs are processed, both in neural activity (Koenig and Lehmann 1996, Pulvermüller 1996, Shapiro et al. 2001) as well as in priming, production, and speech errors (Faye and Cutler 1977, Shapiro and Caramazza 2003, Abrams, Trunk, and Merrill 2007, Mirman et al. 2009, Park et al. 2013). Others have suggested alternative explanations at the root of this phenomenon (Vigliocco et al. 2011). Likewise, other studies have investigated psycholinguistic differences in open (content) versus closed (functional) class words (Biassou et al. 1997) as well as differences in neurological
processing (Neville et al. 1992, Kutas 1997, Osterhout et al. 2002). Still others have examined the root-affix distinction in particular (Jarvella & Meijers 1983).

These strands of literature in the domains of psycho- and neurolinguistics suggest a few possibilities. First of all, the existence of grammatical categories in languages and their observed patterns could reflect innate or universal differences in their treatment. Secondly, grammatical categories may form cohesive classes of words or neighborhoods. Thus, words stored in the lexicon must be categorized in some way according to their classes, and words sharing the same class are treated as sharing a psycholinguistically relevant similarity. One might hypothesize that either of these factors could be possible sources of the phonological asymmetries that linguists have observed. That is, sound changes or phonological rules may spread throughout the lexicon between words sharing any type of similarity, whether it be phonological, semantic, or in sharing a grammatical category.

Despite many observations from the literature on grammatical category processing, there is little evidence to suggest that grammatical categories themselves have innately different means of storage and processing within the lexicon. Additionally, while there are strong neighborhood and priming effects observed among words sharing grammatical categories, much of the evidence points to overlapping syntactic and semantic properties as being the conditioning factors. I argue here that grammatical class phenomena can be attributed to syntactic and prosodic level effects, suggesting that morpheme categories need not be specified in the lexicon to yield the class-specific phonological asymmetries which have been documented.

Lastly, in light of this finding I consider a third hypothesis: Neurological and psychological factors may be relevant at the syntactic, rather than lexical level. A number of studies (Lindblom 1995, Hickok and Poepel 2004, 2007) have shown the activation of two different neural streams
which may be involved in different linguistic tasks. These two streams may be involved with two
types of listening, one at the semantic level (identifying words), and the other at the phonetic level
(identifying sounds). Along with literature considering the role of predictability in speech
perception (Cole & Jakamik 1980, Alyett & Turk 2004), these observations suggest that phonetic
listening is more strongly activated in less predictable environments. This phonetic activation is
then mapped to the syntactic and prosodic structure of a language, resulting in phonological
faithfulness in some positions, but reduction in others. This model accounts for the more sharply
contrasted phonological differences observed in open versus closed classes, affixes versus roots,
and function versus content words, as opposed to the more conservative differences displayed
among nouns, verbs, and adjectives.

This paper will be organized as follows. Section two will review a sample of category-
specific phonological patterns. Section three will review the literature concerning psycho- and
neurolinguistics effects specific to noun-verb differences and will address whether such categories
are active in the lexicon such that they could be involved in conditioning phonological effects.
Section four will focus on the function versus content morpheme distinction, first reviewing the
literature on observed differences in psychological treatment, and then considering lexicon-based
hypotheses. Here I will also consider the dual stream model of speech perception and how it might
influence category-specific effects at the syntactic level. Section five will discuss the proposed
model and consider specific questions about its functioning and other remaining issues, while
section six will conclude the paper.

2. A Brief Typology of Category-Specific Phonological Patterns

A wide selection of studies have documented and analyzed examples of phonological
patterns which appear to have been influenced by grammatical categories of various types. These
examples, however, come from at least two distinct strands of literature: diachronic linguistics, where a primary goal is explaining the motivation for sound changes, and synchronic phonology, where the focus is typically on modeling the patterns according to rules in an active grammar. From these two bodies of literature come three basic types of class-specific patterns. The first of these involves regular sound changes, where the synchronic result may be phonological subinventories, with different allowable contrasts for different grammatical classes. The second strand of literature concerns lexical diffusion, where sound changes have been reported which spread through different classes at different rates. Finally, there are cases of phonological rules which may only be activated in words or morphemes of specific classes. Despite these differences, all these cases are unified in their exhibition of phonological asymmetries specific to different categories.

2.1. Grammatical Categories and Regular Sound Change

The first set of examples involve cases of regular sound change which involve grammatical categories as environments. Much as when a sound change takes a phonological environment as a conditioning factor (e.g., X becomes Y word finally, intervocally, when before sound Z, etc.), some have argued that sound changes can have grammatical classes as environments (e.g., when in verbs, affixes, function words, etc.). This concept is at odds with traditional principles of sound change such as those proposed by the Neogrammarians, such that sound change is a “purely phonetic process” occurring under “strictly phonetic conditions” (Bloomfield 1933:364, concerning the Neogrammarians). Despite this, there are in fact examples of sound changes which appear in a synchronic analysis to have followed such conditions.

For example, Crowley (1997) presents a case of a sound change with “grammatical conditioning” in the Northern Paamese language of Vanuatu. Generally speaking, the language
underwent a sound change in which historic *l was lost in the phonological environment preceding a non-high vowel such as /e/ or /o/. However, this rule was blocked in verbs; verb forms such as *leheie ‘s/he pulled it’ and *loho ‘s/he ran,’ retained their initial /l/, occurring as /lehei/ and /loh/ in modern Paamese as opposed to /ehei/ or /oh/ which would be expected for a noun having the same historic form.

Examples of this sort may also result in class-specific subphonologies that affect the allowable contrasts in different classes. A well known example of this comes from English in the development of the voiced interdental fricative, /ð/. While /ð/ developed in medial position in a phonological motivated environment, intervocally, in word initial position the original /θ/ became voiced /ð/ only in certain more functional categories such as pronouns (‘thou,’ ‘thy’), articles (‘the’), demonstratives (‘this,’ ‘that’), and functional adverbs (‘then,’ ‘though’). In content word categories, such as nouns (‘thing,’), verbs (‘think’), adverbs (‘thick’), prepositions (‘through’) and content adverbs (‘thickly’) the sound change did not apply.

A fairly similar but more complex case occurred in many Athabascan languages in the development of voiced fricatives from voiceless ones. In the Hän Athabascan language of eastern Alaska, voiced fricatives developed in syllable onset position in postpositions, adjectives, compounded noun and verb stems, and verb prefixes in the conjugation and qualifier zones of the verb template, regardless of the phonological environment (Manker 2015). Voiceless fricatives remain in most other word and morpheme categories, although in nouns and verbs these segments alternate in voicing depending on the voicing status of surrounding segments. In any case, the voicing distinction is arguably not contrastive but is predictable based on the phonological and grammatical category environment. In Koyukon, another Athabascan language in western Alaska,
a similar pattern emerges but involves slightly different classes; for example, fricatives in verb stems remain voiced regardless of the phonological environment.

Other examples of class-specific subphonologies abound in the literature concerning synchronic phonological patterns. Smith (2011) catalogues a number of these patterns which occur in different word classes such as nouns, verbs, and adjectives. As she observes, these category-specific effects are overwhelmingly prosodic in nature. For example, in Spanish, stress is contrastive in noun stems and may occur in antepenultimate, penultimate, or final syllable position. On the other hand, stress is not contrastive for verbs but its placement is determined by the inflectional affix. A parallel case occurs in Tokyo Japanese, but concerns the placement of pitch accent, which is once again contrastive in only nouns. In Mbabaram, a Paman language of Australia, long vowels only occur in nouns but never occur in verbs. Smith also finds a number of word minimality requirements that are class-specific. In Chuukese, a Micronesian language, nouns are subject to a bimoraic requirement from which verbs are exempt. Thus a surface contrast between CV:C and CVC syllables exists in verbs which is not available for nouns, since *CVC does not meet the nominal bimoraic requirement.

A number of linguists have analyzed what prove to be even more profound phonological differences in roots (usually content morphemes) and affixes (usually functional morphemes) with a far more numerous set of examples to draw from. In general, phonological inventories available to affixes are subsets of those available to roots, having a smaller number of contrasts available (Willerman 1994). These affix subphonologies typically consist of less marked sounds. For example, in Classical Arabic, affixes do not contain pharyngeal consonants but they are permitted in roots (McCarthy 2005). In Lushootseed, a Salish language, ejective sounds are prohibited in affixes, but commonly occur in roots (Bates et al. 1994). Vowel reduction can also be specific to
affixes. In Hän (Athabascan), the phone [ə] only occurs in affixes; additionally, long vowels are not permitted in affixes (Manker 2012). For many other examples of root-affix phonological asymmetries, refer to Beckman (1997, 1998) and Urbanczyk (2011).

2.2. Grammatical Categories and Lexical Diffusion

While the Neogrammarian hypothesis claims that sound changes are exceptionless and affect all instances of the relevant sounds in a given environment, the theory of lexical diffusion proposes the following alternative: that a sound change may begin in a few words and then gradually spread throughout the lexicon (Wang 1969). That is to say, words, and not abstracted phonemes, are eligible vessels for the spread of a sound change. However, this diffusion is not entirely randomly, and may be facilitated by similarities shared among words. For instance, Bybee (2001) argues that frequency, in addition to shared phonological characteristics, can influence a sound change’s progression. Yaeger-Dror (1996) suggests the influence of lexical classes on the spread of a sound change, where shared semantic characteristics facilitate lexical diffusion.

Others have suggested that grammatical categories could likewise act as influencing factors in sound change, such that words or morphemes of the same category would be considered more similar or more cohesive in some way, facilitating the spread of phonological information. A few studies have shown synchronic patterns of this type. Phillips (2006), using corpus data, finds an effect of word class in the sound change involving the laxing of long /u:/ to /u/ from Middle English to Early Modern English, showing that verbs tended to adopt the change before nouns. She finds a similar effect on the spread of the diphthongized variant of /ij/ in Dutch. Likewise, Hoffman (2010) finds an effect of word class on the velarization of /n/ in Spanish while Brown & Torres Cacoullos (2003) show a similar effect in word final /s/ debuccalization, yet Brown (2013) offers a competing analysis of the apparent asymmetry.
Another case of a diffusive type sound change with a class condition comes from Donohue (2005) which he analyzes as a change in which “the only delimiting factor… is the syntactic category of lexical items” (p. 429). In the Palu’e language of Indonesia, a process of intervocalic stop voicing is active only in genitive clitics. Thus, *ku becomes /gu/ and *ta becomes /de/. In this case, the category is not as much syntactic (noun, verb, etc.) but morphological (affix, root, clitic). Donohue interprets the development as a case of lexical diffusion, where the shared clitic class (and possibly the shared genitive function) is licensing the spread.

2.3. Grammatical Categories and Active Phonological Rules

A number of active phonological rules also seem susceptible to the influence of grammatical categories, such that these rules either fail to apply in words of certain categories, or are applied differently. For example, Shiraishi (2004) describes an example from Nivkh which involves an alternation between fortis consonants--- including aspirated stops and voiceless fricatives, and lenis consonants--- including plain stops and voiced fricatives. Compounded verb stems undergo a process of consonant fortition while compounded noun stems are exempt from this rule. In Hän (Athabascan), stem initial fricatives alternate in voicing depending on the surrounding phonological environment, but prefix initial fricatives are static and retain their underlying voicing status regardless of the environment (Manker 2015). McCarthy (2005) shows a case from Moroccan Arabic where consonant clusters must be broken up with schwa insertion; however, the insertion rule is different for nouns and verbs. Ultimately, given the much stronger relationship of these changes with morphology, these cases can often be interpreted in terms of base-identity effects (Shiraishi 2004, Cable 2004), differences in class-specific affixation patterns (McCarthy 2005), or even that the phonological rule could be specific to a morphological process.
As a result, such cases are considered of lesser importance to the present paper, since they are less suggestive of category-specific cognitive differences.

3. Cognitive Differences and Noun-Verb Phonological Assymetries

The patterns outlined in section two show many cases of phonological asymmetries among different grammatical categories. What, then, are the underlying factors motivating this outcome? One possibility which has been hinted at in much of the literature is that there are cognitive reasons that might explain this pattern. Two specific hypotheses are outlined below.

One possibility is that there are innate or universal phonological constraints governing grammatical categories; that is, categories have optimal or ideal forms, and sound changes are bound to develop over time in order for words of different classes to drift towards these ideal forms. This could account for the examples of category-specific regular sound change that can result in phonological subinventories for different categories. While studies in synchronic phonology typically do not focus on the ultimate source of surface patterns, some analyses in this branch of literature could indirectly suggest this as a possibility. For example, studies in the literature on root-affix phonological asymmetries have developed the notion of a root faithfulness constraint within the grammar (McCarthy 1995, Steriade 1995, Beckman 1998). In the framework of optimality theory, such constraints are usually considered part of the universal grammar (UG), but the ranking respective to other constraints differs from language to language (Prince & Smolensky 1993). Thus, following these analyses, the fundamental motivation for category-specific phonological asymmetries would seem to be a root-faithfulness constraint in the UG, and not more basic factors.
In the literature on phonology and part of speech, Smith (2001, 2011) has also proposed a set universal constraints following an expected hierarchy, with phonological faithfulness to nouns being ranked the highest (Nouns > Adjectives > Verbs). Such constraints are used to model the observed patterns in the synchronic grammar, and again, there is less focus on the underlying motivations of this apparent hierarchy. Smith (2012) does address this point briefly, speculating it could be due to a “bias in language acquisition,” citing literature showing differences in part of speech in both acquisition and neurolinguistic studies.

The second possibility is that grammatical categories form neighborhoods that influence the spread of a sound change throughout the lexicon. Just as shared phonological characteristics might facilitate the spread of a change, a shared grammatical category could act in a similar manner. This could account for the examples of category-specific imbalances in lexical diffusion. Although not directly stated, such a hypothesis could be implied from Phillips’ (2006) account of the effect of word categories on lexical diffusion.

This section will consider the literature from a parallel branch of research which investigates how nouns and verbs function differently within the human brain. One branch of this research in psycholinguistics investigates the question of what human behavior tells us about the structure of the mind and its cognitive processes. Studies in this field focus on priming effects, word recognition, and speech errors, yielding in some cases category-specific patterns and behavior. The other branch is in neurolinguistics, which focuses on the physical parts of the brain involved in the storage and processing of different words, as well as the electrophysiological response of the brain.

What then are the necessary criteria for supporting the possibility that there are universal constraints on optimal phonologies for different grammatical categories, or that those categories
form neighborhoods to facilitate the diffusion of sound change? A first necessary criteria for allowing either of these hypotheses is that grammatical categories must be present and accessible within the lexicon. If the phonology cannot access grammatical categories in the lexicon, we should not expect to find true examples of class conditioned sound change or lexical diffusion. Following the observations of Vigliocco et al. (2011), I will show that the much of the literature on category differences does not provide support for the activation of these categories within the lexicon, but rather at the post-lexical or syntactic level.

Secondly, studies in category-specific cognitive differences must also distinguish semantics from word categories. A common observation (Kemmerer & Eggleston 2010, Vigliocco et al. 2011) is that prototypical nouns refer to objects, while prototypical verbs refer to actions or events. However, this is not always the case (e.g., “the run,” refers to an event of running and not an object). Thus, results showing differences in nouns and verbs which do not separate these two variables may be demonstrating cognitive differences in how semantic factors are treated rather than grammatical category. By addressing the question of whether grammatical categories are active within the lexicon, I will simultaneously provide evidence against both proposed hypotheses: that there are universal constraints governing category-specific phonology or that categories form neighborhood by which sound changes can diffuse throughout the lexicon. A review of the literature shows no evidence for different cognitive treatments of different categories within the lexicon when other factors, such as syntax and semantics, are considered.

3.1. Psycholinguistic Behavioral Differences in Nouns and Verbs

While linguists have for long considered nouns and verbs to be basic categories in languages, existing as a fundamental distinction in nearly all languages (Greenberg 1968), only
more recently have researchers studied differences in their psycholinguistic treatment. These studies have focused on patterns of speech errors, aphasia, and priming effects.

Faye & Cutler (1977) was an early study which detailed something of a neighborhood effect for word categories in the occurrence of the speech error known as a malapropism. Malapropisms are errors in which the speaker unintentionally replaces a target word with one that is phonologically similar but semantically unrelated. Their results showed that in 99% of cases, a word of the same grammatical class as the target word was chosen. This finding suggests what could be a strong cohesive effect among words of the same categories, such that words sharing the same categories would be considered more similar. Here it is clear though that words were in a syntactic, grammatically constrained context, and thus this is not an indicator that the categories were functioning as neighborhoods within the lexicon.

More recent studies in priming tasks have shown similar results. Abrams, Trunk, and Merrill (2007) investigated the tip of the tongue phenomena whereby speakers feel they have a target word in mind but cannot recall its phonological form. In their experiment, subjects were asked to answer short general knowledge questions. In the event the subject could not remember the answer, a phonologically similar prime was supplied. Their results showed that a word with a similar phonological form but of a different grammatical class was a better prime for the target word than words with of the same grammatical class. This result can be interpreted such that words sharing grammatical classes share more similarities, and thus provide more competition for the target word, and interferes rather than aids in its retrieval. As with Fay & Cutler (1977), the target words used in Abrams, Trunk, & Merrill (2007) were presumably in a syntactically constrained context which resulted in the elimination of phonological similar but grammatical incongruent candidate words (they do not give examples of the questions that were used, however).
Additionally, their study shows the strength of phonological priming and that priming still occurs among words of different categories, suggesting the primary characteristic indicating similarity is phonology and not word category.

Strand, Simenstad, & Cooperman (2014) investigated how the grammatical density of a word affects its recognition in the context of a sentence. Grammatical density is a calculation of how many phonologically similar words a particular word has within its grammatical class. Subjects were asked to identify a word after it was spoken, both in a sentence and in isolation. In sentences, where words had a “grammatically constrained” context---meaning their part of speech was determined by surrounding words---a lower grammatical density resulted in more successful recognition of the word. In an unconstrained context, grammatical density did not matter. Strand et al. analyzed these results as evidence that speakers use knowledge of part of speech to eliminate incongruous word candidates. A second experiment measured response time in determining whether the words were real English words or not; additionally, pseudo-words were included. Grammatical density can be calculated for pseudo-words since their category can be determined in a sentence context, and pseudo-words still bear phonological similarity to real words. However, in lacking semantic content, they avoid any bias in activating semantically similar words. As expected, it was found that subjects had a quicker response time in recognizing both words and non-words in a grammatically constrained context. Ultimately, Strand et al. propose adding an additional grammatical class layer to the TRACE model of word recognition above the currently accepted word, phoneme, and feature levels. This would account for the effect of eliminating words of incongruent word classes in a constrained context, and may also suggest an additional degree of separation between words of different grammatical classes. However, once again, the results only show sensitivity to grammatical categories in a syntactically constrained context, and
it would not follow that a grammatical category layer would need to be present in the architecture of the word recognition model when processing words in isolation.

Mirman et al. (2009) conducted two word recognition experiments, the first of which presented listeners with an auditory lexical decision task in which they had to press a key once they recognized a word in isolation as being a real English word or not, while the second involved an eye-tracking paradigm where subjects were asked to look at the image that matched the word they heard. The stimuli included words with a same class homonym (‘deck’ vs. ‘deck,’ both nouns), words with different class homonyms (‘bark’ n. vs. ‘bark’ v.), and words with no homonyms. The results showed that words with a same class homonym took subjects longer to recognize than words with different class homonyms. Like Abrams, Trunk, and Merrill (2007), these results suggest that there is more competition for recognition among words sharing not only a phonological form, but a grammatical category. However, Mirman et al.’s (2009) experiment differs crucially from those of the other psycholinguistics studies reviewed here, because theirs focused on the recognition of words in isolation rather than in a syntactically constrained context, where other studies have shown strong evidence for a similar patterning of words sharing grammatical categories. Despite this, Mirman et al. (2009) did not separate the overlapping semantic variables, and thus their homonym pairs different not only in grammatical categories, but in semantic classes. ‘Bark (of a tree)’ is both a noun and an object, while ‘bark’ is both a verb and an action. Vigliocco et al. (2011) suggest this is a common oversight in studies making claims about the cognitive processes involved in grammatical categories. Vigliocco et al. (2006) isolates these variables in a neurological study (see 3.2.) showing the differences in grammatical categories are neutralized when semantics are controlled for.
3.2. Neurolinguistic Differences in Nouns and Verbs

Additionally, a number of studies have investigated neurological differences in syntactic categories, or parts of speech, such as nouns and verbs. This branch of literature includes brain lesion and aphasia studies, imaging of neurological activity, and analysis of electrophysiological energy patterns. As with the findings in psycholinguistics, research in neurolinguistics consistently shows differences in how different grammatical categories, primarily nouns and verbs, are processed by the brain.

Another topic of investigation has been the behavior of brain lesion patients, who have typically suffered injuries resulting in speech deficits. Often these lesions are localized to specific areas of the brain, and subjects may only have deficits in some aspects of speech but not others. One curious phenomenon is that some patients have significantly more difficulty processing (reading, writing, understanding or producing) verbs than nouns, while others have difficulty producing nouns but not verbs (Silveri, Perri, & Cappa 2003).

One obvious complication in many of these aphasia studies is not separating grammatical class and semantic variables, and the nouns are overwhelmingly objects while the verbs are overwhelmingly actions. Park et al. (2013) observed patients with Broca’s aphasia who had more difficulty producing verbs than nouns. Their study investigated the production of verbs that had a conceptual or phonological relationship to a noun. For example, to sweep shares the conceptual relationship of instrumentality with the noun broom, and to hammer is phonologically (and conceptually) related to its noun homonym, hammer. The results of their study showed the Broca’s patients were significantly more successful in producing verbs with either a phonological or semantic relationship to another noun. This study suggests less of a clean split between different grammatical categories. Verbs with close relationships to objects were easier to retrieve, which
indicates a deficit in producing words of a particular semantic class rather than a grammatical class.

Building on the observations of aphasia patients which category-specific deficits, Shapiro et al. (2001) conducted a study to determine the parts of the brain which are active in processing of different parts of speech. One primary question addressed in this study is whether different parts of the brain are used for different words based on actual grammatical class, or semantic characteristics. This is due to the fact that canonical nouns tend to represent objects while verbs typically represent actions. Shapiro et al. hypothesized that the left frontal cortex was involved in processing verbs and used a technique of repetitive transcranial magnetism simulation (rTMS) to suppress brain activity in the left frontal cortex. Subjects were given a task of producing inflected noun and verb forms. It was found that the suppression of the left frontal cortex led to a significant increase in response latencies for verbs with no effect on noun inflection. A second experiment involved the inflection of pseudowords. Pseudowords lack semantic content and should therefore not act different based on semantic properties such as actions or objects, yet verb pseudowords showed the same effect as real words when the left frontal cortical was suppressed. Shapiro et al. (2001) interpreted the findings as evidence that different portions of the brain are recruited for the production of words in different grammatical categories, and that this difference was not the result of semantic properties of the words. Shapiro (2003) built on these results by finding that aphasics with noun-specific deficits had difficulty inflecting noun pseudowords, while aphasics with verb-specific deficits showed difficulty inflecting verb pseudowords.

While Shapiro et al. (2001, 2003) show strong evidence for neurological differences in the way nouns and verbs are processed, even involving words in isolation, they elicited inflected forms and not entries that would appear in the lexicon. Thus the findings suggest neurological differences
in noun and verb inflection, not nouns and verbs themselves. This is supported by another observation of Vigliocco et al. (2011) which notes the difference in processing complexity that nouns and verbs have, such that verbs involve more complex processing in having to coordinate between subjects and other possible arguments. Additionally, Shapiro’s first experiment does not remove the semantic variables from the word comparison, resulting in the action vs. object overlap. Their second experiment does attempt to address this by using pseudowords. However, Vigliocco et al. (2011) also warn against the assumption that pseudowords contain no semantic information. Their own study found that subjects overwhelmingly assigned object meanings to pseudo-nouns and action meanings to pseudo-verbs. Regardless, the use of inflected forms suggests nothing about the status of grammatical categories in the lexicon, so it would not suggest the possibility of phonological sensitivity to these categories.

Lastly, a number of studies have examined patterns of electrocortical energy produced when subjects hear words of different categories. These studies measure fluctuations in the electric field topography of the scalp by measuring the voltage output at different locations. These fluctuations are indicative of a change from one neural process to another (Koenig & Lehmann 1996). This enables the measurement of event related potentials (ERPs) which are patterns of electrical activity in response to a particular stimulus. Koenig and Lehmann (1996) conducted a study showing differences in ERPs for nouns and verbs heard in isolation. In particular, there was a difference in electrical activity in the 116-172 msec post-stimulus microstate, which suggested the activity in the brain “differed in location and/or orientation” (pg. 117).

Pulvermüller et al. (1996) conducted a similar study examining differences in electrocortical energy for nouns and verbs but measured the activity not only at different locations on the scalp but at different frequencies. Nouns and verbs showed significant differences in the
energy output at a frequency of 30 Hz: nouns displayed stronger activity at 30 Hz at locations over the visual cortices, while verbs displayed stronger activity in locations over the motor cortices. This study showed that nouns and verbs not only utilize different parts of the brain, but that these parts correspond to areas used for parallel non-linguistic uses, such as vision and movement. This also aligns with the generality that nouns are objects and verbs are actions, and suggests the results of Koenig & Lehmann (1996) could be due to the confound of grammatical class and semantics.

Pulvermüller’s findings fit well with Vigliocco’s study of Italian (2006). Vigliocco used PET imaging to identify the portions of the brain active when listening to words. Vigliocco et al. (2006) chose only event nouns and verbs and further distinguished between sensory and motor events (for example, motion words *corsa* n. ‘run,’ vs. *galoppa* ‘s/he gallops’ or sensory words like *solletico* ‘tickle’). Sensory words of both categories activated the anterior temporal cortex while motion words activated the primary motor cortex, areas also involved in processing non-linguistic sensory and motor information. The study found no activations unique to nouns and verbs, suggesting noun and verb comprehension “engage[s] a common neural system” (Vigliocco et al. 2011, pg. 421).

In reviewing the psycho- and neurolinguistic literature, studies on speech errors, priming, word recognition, the behavior of aphasia patients, response to localized rTMS, and differences in electrocortical energy suggest profound differences in the cognitive processing of grammatical categories like nouns and verbs. However, these findings show no evidence that these syntactic categories are active or accessible within the lexicon. Rather, distinctions between noun and verb categories are activated at the morphological or syntactic levels. Secondly, semantic classes such as action and object words, as well as sensory or motor words, display both psycholinguistic and
neurological differences, even in isolated words, which suggests these characteristics are active in the lexicon. The frequent overlap between semantic class and grammatical category is a confound resulting in the appearance of category-specific behavior in the lexicon. Thus, since the current literature does not provide support for active syntactic categories in the lexicon, categories can neither have universal phonological constraints nor act as neighborhoods by which sound change diffusion is facilitated.

4. Cognitive Difference in Function and Content Morphemes

Descriptions of category-specific phonology for open vs. closed morphemes, functional vs. content words or roots vs. affixes (Steriade 1995, Beckman 1998) show much more profound differences than the typically suprasegmental differences sometimes observed in nouns and verbs (Smith 2011). Here we will observe a weaker boundary between morphology and syntax and consider affixes (especially inflection), function words, and closed class words together as a group, compared to roots, content, and open class words. This better captures the similarity in that functional morphemes tend to undergo more extreme reduction, while content morphemes tend to resist this phonological erosion. While syntactic categories such as nouns and verbs show no clear cognitive differences within the lexicon, perhaps the more profound differences between open and closed class morphemes can be captured by lexicon-based cognitive differences.

In section 4.1. I will first review the striking grammatical polarities that functional and content morphemes exhibit. In light of these differences, in 4.2. I will consider three hypotheses which may explain the duality of these classes in terms of cognitive differences. The first two are the same considered for noun-verb differences---that category-specific phonology is the result of universal constraints, or that diffusion is expected to occur within these categories at a faster rate due to sharing a particular feature. The third hypothesis is that cognitive differences in these
classes are not innate but emergent, and are active in the syntax rather than the lexicon. While profound differences are once again observed in psycholinguistic and neurological studies, there is still a lack of clear evidence for these categories being active within the lexicon when syntax and semantics are considered. Rather, a review of literature on the dual stream model of speech perception in section 4.4. lends support to the third hypothesis.

4.1. Grammatical Differences in Function and Content Morphemes

4.1.1. Phonological Differences. As reviewed in section 2, profound phonological differences are displayed between content and functional class morphemes. It is crucial to point out that these differences are much more profound and much more common than those displayed among open class categories such as nouns and verbs. Functional morpheme class phonology is virtually always reductive in nature, and often subinventories form containing only a portion of the constraints available for open class morphemes. Lastly, it should be noted that these differences are much more likely to be segmental rather than the primarily suprasegmental differences observed between nouns and verbs.

4.1.2. Syntactic Differences. Content and functional morphemes differ syntactically primarily in that closed class words display a much higher degree of predictability in a syntactic context. Functional morphemes include subclasses such as inflectional affixes, determiners, prepositions, and pronouns. Any of these classes has a fairly limited number of possible morphemes, and depending on the language may include between only a few and a hundred or so candidates. New members of these classes cannot easily be borrowed from other languages or created. By contrast, a language’s nouns and verbs will number well in the thousands. In the presence of a sentence, the existence of certain elements will often determine the presence of others based on the phrase structure rules of a given language; For example, the presence of a
determiner in English implies a noun must be soon to follow; the presence of a noun indicates a strong likelihood a determiner occurred somewhere not long before. In the case of the latter, the number of possible determiners in English is quite small. Thus, the presence of a noun makes the presence and identity of determiner much more predictability. A final crucial point to consider here is that because such predictability is tied to the number of available morphemes in different classes, and the phrase structure rules of a given language, this characteristic of functional morphemes must be emergent, or learned from a speaker’s observation of his or her language, rather than being an innate or universal characteristic.

4.1.3. Semantic Differences. Much with the confound of semantics and grammatical class with nouns and verbs, content and functional morphemes overlap almost perfectly with different semantic classes. Functional morphemes are consider closed, and not open to new members, partially due to the fact that they are the grammatical morphemes of a language, and there are a relatively limited number of grammatical concepts that can be expressed (compared to the number of objects or actions in the world). Thus, closed class functional morphemes are always grammatical and highly abstract; they rarely if ever have real word referents. Content morphemes can be quite abstract, but are never functional, and always have content-based referents. Thus, there is almost always a perfect overlap of unique semantic characteristics associated with functional and content morphemes respectively. Like the differences between actions and objects, there may also be innate differences between how functional and content semantics are processed in the mind. This added variable cannot easily be controlled for, and as a result cognitive differences in content and functional morphemes in isolation may tell us very little about what the source of the difference is. However, we can still see evidence of emergent cognitive differences
in the syntax if greater differences occur between the two classes in the context of sentences rather than when words of these classes occur in isolation.

4.2. Cognitive Differences in Functional and Content Morphemes

4.2.1. Aphasia studies

Researchers for long have observed patients with Broca’s aphasia who are agrammatic, displaying a sort of telegraphic speech consisting of content words but few or no function words. For example, Gardner and Zurif (1975) observed Broca’s patients who showed a higher error rate when reading grammatical particles (such as ‘or,’ ‘at,’ etc.) compared to picturable nouns. Andreewsky & Serron’s (1975) study presented French words that have both functional and content homographs (mais as either the conjunction ‘but,’ [mɛ] or the noun ‘corn’ [mais]) to Broca’s patients, finding they were usually pronounced as if open class words. Such observations led to Bradley’s (1978) model of a dual lexicon, with a separate lexicon for roots and another lexicon for affixes. The existence of two separate lexicons could fit perfectly with either a hypothesis of universal phonological constraints unique to each lexicon, or that diffusion could occur within one of these lexicons but not the other.

Later studies, however, questioned the need for two separate lexicons, and rather proposed dual processing of open and closed class words. Bradley’s (1978) findings were based on the fact that normal subjects showed more errors in producing low frequency content words, but that frequency did not have an effect on the production of function words. By contrast, for aphasics, word frequency was relevant for the production of both open and closed class words. Bradley interpreted the results as indicating that for normal speakers, the insensitivity to frequency for closed class words implied a separate lexicon. However, Biassou et al. (1997) and others were
unable to replicate Bradley’s (1978) results. Biassou et al. (1997) argued that the results were consistent with a single lexicon, and that two routes for processing were involved after lexical retrieval.

A model of dual processing, rather than dual lexicons, may not imply a different cognitive treatment of open and closed class morphemes at the lexical level, and such differences in processing could take place at the syntactic level. On the one hand, Nespoulous et al. (1988) found that agrammatic aphasics had no trouble producing closed class function words in isolation, while their primary difficulty involved producing these words in the context of sentences. On the other hand, Biassou et al.’s (1997) study found similar rates of phonological errors for closed class function words whether in isolation or sentence context. However, it should be noted that the difference in error production for closed vs. open words in isolation was only marginally significant ($p < .1$) compared to strong significance ($p < .001$) for the difference in error rate of these words in a sentence context. This could suggest the processing distinguishing content and function words is magnified in the syntactic context, though the inconclusive results still leave open the possibility for lexicon-based distinctions.

4.2.2. Differences in Listener’s Attention to Content and Functional Morphemes

Another facet of the cognitive differences between content and functional morphemes is their prominence in the minds of listeners, that is, how much attention is being paid to morphemes of these different classes. Several studies show a common theme indicating greater psycholinguistic prominence of content morphemes, with less attention given to functional morphemes.
Jarvella & Meijers (1983) looked specifically at differences in stem and affix processing in Dutch. They found that subjects displayed a shorter response time in identifying similarity of words sharing a stem rather than sharing the same inflectional form (for example, tense). They argue that words are accessed based on their stems, with affixes being accessed only secondarily. Here, it should be noted that this difference is perceived in inflected forms, and does not reflect differences in these morphemes in the lexicon.

Several studies within the literature on language acquisition discuss the lateness of learning inflectional affixes in particular (Polišenká 2010, Penke 2012) and explain this in terms of the salience of inflectional affixes. In particular, affixes are shown to have lower phonological salience and in English tend to be shorter than roots and may only exist as a single consonant in a coda cluster. Of course, the suggestion here is that the lower phonological salience is due to affixes’ phonologically reduced forms, and not necessarily a characteristic of being a member of a functional morpheme class. Nevertheless, the two variables here, the grammatical category and phonological reduction, are impossible to separate from the acquisition observations alone. Once again, differences in phonological salience are only shown to arise in inflected word forms and not within the lexicon.

Other studies have looked at differences in event related potentials (ERPs) for open and closed class morphemes. Neville at al. (1992) found different electrophysiological responses for open and closed class words in sentence contexts. Closed class functional words displayed an earlier N280 component (a negative potential about 280 ms after a stimulus) with a greater amplitude. This was strongest on the anterior temporal and temporal locations in the left hemisphere. Open class content words had a stronger N350 component, located on the posterior
regions of both hemispheres. She concludes that such results contribute to an understanding of the “different functional subsystems” utilized when processing words of different classes.

Figure 1: Averaged ERPs measured at the left frontal region for five lexical classes, from Kutas (1997)

Kutas (1997) discusses some of Neville et al.’s (1992) findings which display a stronger N400 component (N350 is considered a member of the larger N400 component class) for open class words. While Neville had noticed some relationship between N400, frequency, anomalous or unexpected semantic information, Kutas goes as far as saying that stronger N400 components were indicative of “semantic expectancy and its consequences for online processing and not lexical class” (pg. 385). Additionally, Kutas (1997) showed that N400 was typically diminished in amplitude near the end of sentences, where there was more syntactic and semantic context. Furthermore, multiple repetitions of open class content words could elicit a reduced N400
component. As for N280, Kutas showed that its peak latency correlated more with word length and word frequency and the higher frequency and shorter length of closed class function words was the cause of the stronger N280 component. Thus, Kutas concluded that, given these different latencies (about 280 ms for closed class words and 330 ms for open class words) for what she calls lexical processing negativity, as well as semantic expectancy, that the scalp distributions for open and closed class words are “remarkably similar” (pg. 387). Overall, this suggests differences in the amount of attention at the syntactic and semantic, rather than lexical level, and thus differences in processing needed for content and functional morphemes is more accurately a correlate of the semantic and syntactic predictability of those words.

4.3. Content and Functional Morphemes and Arguments for Lexicon-Based Differences

Section 4.2.1 showed radically different treatment of function and content morphemes by aphasics, such that some patients were completely incapable of producing function words. This led some (Bradley 1978) to propose dual lexicons for open and close class words. Such an analysis would harmonize well with a hypothesis that either universal phonological constraints or category-specific lexical diffusion is the cause of the phonological asymmetries displayed by these two classes. However, Biassou’s (1997) analysis suggested a difference in processing. While she found that aphasics had similar difficulty in producing function words in isolation (as opposed to the findings of Nespoulous et al. 1988), the differences were more significant with function words in a syntactic context. Some word level processing effects may still take place with words in isolation as closed class words are semantically functional and may still weakly activate portions of the brain engaged in grammar. However, given the complete overlap between category (open vs. closed) and semantic class (content vs. function) it is impossible to separate these variables.
In addition to this inconclusive evidence, section 4.2.2. presented cognitive differences displayed by normal functioning speakers, showing a difference in the perceptual salience between the two morpheme categories. All of these studies looked at differences in functional and content morphemes either in inflected words or in the context of sentences, providing no evidence of lexicon-based differences. Furthermore, Kutas (1997) showed direct evidence that the differences in neurological behavior, even though observed in the context of sentences, was due to contextual predictability (something which can only arise with context, i.e., at the syntactic level) and not grammatical category differences. The following section will consider dual models of speech perception that have been proposed that align well the behavior of content and functional morphemes, and which may ultimately account for the profound phonological differences that characterize these two classes.

4.4. The Dual Stream Model

Our discussion so far has shown general patterns of duality for content and functional morphemes. Such duality is manifested in phonology, whereby functional morphemes display a high degree of phonological reduction and prominence avoidance, whereby content morphemes display more faithfulness and resistance to erosion. Duality is also seen in syntax---functional morphemes are more predictable based on the syntactic context of a sentence. Additionally, there are asymmetries in semantics in the split between functional and content meanings. The literature has shown polar treatment for content and functional morphemes in the behavior of aphasia patients, cognitive processing, language acquisition, and neurological patterns.

A separate branch of literature has developed that proposes two different types of listening involved in processing the speech signal. Lindblom (1995) describes a “what” and a “how” mode of listening which may be used in different contexts based on their predictability. The “what”
mode of listening is used to comprehend the intended meaning of perceived utterances. The listener focuses only on identifying the words and stringing them together into meaningful phrases and sentences. The “how” mode of listening involves more attention to phonetic detail---details which are largely superfluous for a basic understanding of the meaning of a sentence with ample context. A listener’s use of the two perceptual modes would reflect the level of attention needed to understand meaning: when context can fill in phonetic details based on top-down information, the “what” mode is used, whereas when less syntactic or semantic context is available, listeners must process more phonetic detail for word recognition and activate the “how” mode. Thus, we expect “how” listening closer to the beginning of utterances or when new information is introduced.

Parallels to Lindblom’s “what” and “how” modes of listening are found in neurolinguistic literature, showing that different neural streams are involved in processing speech. Hickok and Poeppel (2004, 2007) define two streams of processing involved in what they refer to as speech recognition and speech perception. The ventral stream involves the retrieval of lexemes from the acoustic signal, that is, a mapping between the phonetic input and semantics (speech recognition). While early speech processing begins in both sides of the brain in the superior temporal gyrus, the ventral stream involves projection ventro-laterally toward the inferior posterior temporal cortex (bilaterally). Here the sound to semantics processing occurs, interfacing then with various regions of the brain where conceptual information (visual, motor, etc.) is stored. The dorsal stream is involved both in motor planning as well as the sublexical processing of acoustic input (speech perception) such as phoneme identification and rhyming tasks; this dual function suggests a link between production and perception. The dorsal stream diverges from the superior temporal gyrus,
projecting dorso-posteriorally in the region of the posterior Sylvian fissure and progressing finally towards the frontal regions. Such projection is primarily oriented on the left side.

Comparing this model to that of Lindblom (1995), the ventral stream corresponds to the “what” mode while the dorsal stream corresponds to the “how” mode. Thus different types of listening are activated according to the contextual predictability. Other studies have shown a correlation between contextual predictability and phonetic salience, suggesting lower levels of phonetic processing in environments of higher predictability. Cole and Jakamik (1980) propose a series of hypotheses of speech perception suggesting this interface between phonetics and top-down knowledge. They first make the claim that “words are recognized through the interaction of sound and knowledge.” That is to say, listeners pay attention to more than just acoustic cues to determine the words in a sentence, and also consider semantics and syntactic structure. They cite Warren (1970), who first demonstrated a phenomenon known as phoneme restoration whereby speakers essentially hallucinate a missing segment that is expected based on the context. In their experiment, they replaced the [s] sound with a cough in the word *legislature* within the sentence “The state government met with their respective *legislatures* convening in the capital city.” Nineteen out of 20 subjects were not able to recognize the manipulation.

In this example we can imagine that the “how” or dorsal mode of listening was largely inactive and that the sentence was being processed primarily through the ventral “what” stream. As soon the words in the beginning of the sentence, such as “state” and “government,” were recognized, they would begin to prime other semantically related words. The listener would already have these candidates in mind, and upon hearing the first few segments of *legislatures*, perhaps only [lɛʤ] or less, would narrow the possibilities down to some form of the word *legislature*. As soon as the word is recognized, very little detailed phonetic listening, the “how”
or dorsal mode, would be required (only enough to make sure the rest of the word turns out more or less as expected). The complete absence of “how” listening would account for the fact that listeners never actually heard an [s], but still perceived it. Ultimately, the rest of the word was similar enough to prior examples that the listener had heard, and thus the listener’s own phonetic representation may have been activated, and in turn perceived.

In a more recent study, Sivonen et al. (2006) investigated the relationship between predictability and phoneme restoration and how this affected the event-related potential (ERP). In their study, they replaced the initial segment of the final word in sentences with a cough, where the final word was either highly predictable or not. This would test the effect of sentence context on phoneme restoration rather than the preceding portion of the word (as in *legislatures*). They found that N400 component, a marker of semantic expectancy, was not significantly different for the highly predictable words that had been manipulated. This may suggest that speakers had only activated ventral stream listening, and a phonetically imperfect token was deemed similar enough to the expected word, its aberrant phonetic form going entirely unnoticed and failing to trigger dorsal listening.

Lastly, Cole & Jakamik (1980) proposed that “words are accessed from the sounds which begin them.” In one experiment, they demonstrated that words with less predictable first syllables were more quickly recognized (e.g., ‘shampoo,’ as opposed to ‘conceive’). Secondly, a study by Marslen-Wilson & Welsh (1978) showed the speakers were more likely to identify a mispronunciation in the onset of a word rather than at the end (e.g., ‘made’ pronounced with initial [n], as opposed to ‘time’ pronounced with a final [n]). Once again, the environment of higher predictability (end of a word) corresponds to a decrease in perceptual salience of the acoustic signal. In these cases, it seems likely that once the beginning of the word resulted in a predictable
recognition of the entire word, that further phonetic processing was halted and the dorsal stream became inactive.

4.5. The Dual Stream Model and Content and Functional Morpheme Classes

The patterns observed in the literature for open and closed class morphemes logically entail the relevance of the dual stream model of speech perception in their processing: that functional morphemes are primarily processed by the ‘what’ mode or the ventral stream, and that processing of content morphemes may involve at least some level of ‘how’ mode listening through the dorsal stream. Closed class functional morphemes have been shown to be more predictable, primarily in a syntactic context, while open class content morphemes, having a much larger set of members, are less predictable. Contextual predictability was shown to align with ventral “what” listening, so by extension functional morphemes would involve this same type of processing. This is corroborated by observations showing decreased attention given to functional morphemes.

The literature also suggests a connection between listening mode and reduction. While more contextually predictable morphemes are shown to be less perceptually salient, and thus do not activate dorsal processing, higher predictability is also shown to correlate with increased phonetic reduction. The alignment of these characteristics are shown in table 1 below. This lack of processing and perception of the phonetic details of functional morphemes may actually enable reduction to proceed unimpeded. Speakers may, over time, produce large pools of diverse tokens of functional morphemes which not only go unnoticed by listeners, but are recognized and processed quickly and efficiently (see 5.1 for more about the specific mechanisms involved in reduction caused by ventral listening).
<table>
<thead>
<tr>
<th>content words (nouns, verbs, adjectives, etc.)</th>
<th>function morphemes (determiners, auxiliaries, prepositions, pronouns, inflectional affixes, etc.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lower contextual predictability</td>
<td>higher contextual predictability</td>
</tr>
<tr>
<td>less reduction, more faithfulness</td>
<td>greater reduction, less faithfulness</td>
</tr>
<tr>
<td>dorsal stream processing</td>
<td>ventral stream processing</td>
</tr>
</tbody>
</table>

Table 1: Alignment of various phenomena according to morpheme classes

Furthermore, I propose that such dual processing is relevant only at the level of syntax, and as a result is mapped to the syntactic structure. The expectation or requirement of the presence of morphemes of particular classes can only occur within a syntactic context, and without being able to make such judgments, the listener cannot eliminate possible word candidates. Thus predictability is irrelevant in isolated word environments. As a result, I propose there is a grammaticalization of processing types which listeners learn to use upon learning the grammatical rules of their language. As shown below in figure 2, dorsal listening becomes associated with content morphemes slots within the syntax, especially at the beginnings of words or phrases where there is little contextual information built up. Such mapping therefore reduces the amount of real-time calculations needed by the speaker to determine by which mode to process which morphemes.

This grammaticalized association between class and listening mode is suggested by the categorical nature of open vs. closed class phonological assymetries. Considering the English initial /ð/ development, which affected entire word classes, and did not did target individual tokens based on their frequency or contextual predictability, but rather on the generalized predictability of the classes as a whole. Thus, all members of the determiner class, from more frequent members like ‘the,’ as well as less frequent ones like ‘these’ and ‘those,’ underwent the same reduction in the form of voicing lenition.
Figure 2: Syntax to Listening Stream Mapping: Size and color correspond to listening stream in the sentence: “John will read his linguistics books in the library,” red corresponding to regions with increased dorsal listening and blue corresponding to regions with primarily ventral listening. This is in turn mapped to the syntactic structure.

The processing of a functional and highly predictable category such as determiners might play out as follows. If we imagine a sentence such as “John saw the cat,” the listener will first process the string [ʤɑnɔ] and will recognize “John,” a subject noun, and “saw,” a transitive verb. The listener will know a noun phrase is likely to occur in the next few words. Where there is a noun phrase, there must be a noun or pronoun and it is highly likely a determiner will closely precede the noun. The set of possible determiners is fairly small; ‘the’ or ‘a’ are highly likely (and even the choice between these two could be determined by pragmatic considerations), as well as [‘this,’ ‘that,’ ‘my’, ‘your,’ ‘his’, ‘her’, ‘our,’ ‘their’], while others such as ‘these’ and ‘no’ could be ruled out if the noun is singular. In this somewhat extreme case, the listener only needs
to listen for a general CV sequence to confirm the determiner as ‘the’ from the list of likely candidates for the next possible words; further phonetic processing is not needed. Where historically the form may have had a full vowel [i] (as in the citation form), reduction may have occurred as the result of random articulatory variation, but was ignored by the listener, who may have persisted in perceiving the intended phonological form, much as in the phenomenon of phoneme restoration.

5. Discussion

Linguists have described an array of category-specific phonological patterns both for differences in syntactic categories such as nouns and verbs and also for broader categories such as content and functional morphemes. Meanwhile, studies in psycho- and neurolinguistics have shown cognitive differences in the treatment of these categories. While the phonological differences might suggest either universal category-specific phonological constraints or diffusion occurring within categories at faster rates, a review of the literature suggests that none of the effects found point to differential treatment within the lexicon. Such category information remains largely inactive or inaccessible by other linguistic modules until placement into the syntax. Thus these initial hypotheses do not seem to account for category-specific phonologies.

However, the results of a large body of literature have revealed a connection between functional morphemes, low predictability, reduction, and ventral stream ‘what’ processing, in opposition to content morphemes, high predictability, faithfulness, and a relative increase in dorsal stream ‘how’ processing. I have further argued that listening modes have become mapped to syntactic positions, with more activation of the dorsal mode in content syntactic positions, and little or no activation of this mode in functional class positions. While differences in listening modes seem likely to have an effect on category-specific sound change, several issues remain
concerning the exact mechanisms and applicability of the dual stream model to category-specific phonological effects.

5.1. The Role of the Speaker and the Listener

One detail concerning the mechanisms of the dual stream model which remains unclear is the role of the speaker in relation to that of the listener in bringing about category-specific reduction. Is it necessary that both the speaker and the listener adopt contrastive but cooperative roles, or can an account be proposed in which only either the speaker or listener are active in bringing about reductive sound change?

Lindblom (1995) also proposes a role for the speaker in shaping the speech signal for the purpose of efficient communication. When new information is introduced, particularly when context does not aid the listener in recognizing words, the speaker may hyper-articulate in anticipation of the needs of the listener. For more contextually predictable words, there is less concern for articulatory precision and reduced forms of these words are introduced into the pool of variants. In this way, the speaker keeps “a running estimate of the listener’s needs” (pg. 8), which is balanced with the drive for articulatory economy.

This argument is echoed in later studies such as Alyett & Turk (2004) which considered the phenomenon of prosodic prominence. Certain prosodic positions are may be marked with greater prominence in the form of longer duration, increased intensity, or higher pitch, in order to maintain a similar amount of meaningful information throughout the signal. Alyett & Turk note the relationship between prominence and predictability, arguing that “prosodic prominence is a linguistic means of achieving smooth signal redundancy. Prosodic prominence increases… with
unpredictable sections of speech…” (pg. 31), ultimately citing the speaker as the initiator of these changes.

So far, only the role of the listener in category-specific reduction has been considered in the present account. However, one possibility is that the speaker and the listener must align in both production and perception for communicative competence---the result being an unintended phonologization of reduction for functional morphemes. While the speaker shapes the signal in order to give prominence to syntactic positions of lower predictability (content words) and hypoarticulates predictable positions (functional morphemes), the listener may need to be aligned to this model for optimal word recognition. For example, while it may seem perceptually advantageous (though more mentally taxing) for the listener to process all words at a phonetically detailed level, a more word-based or phonetically generalized listening mechanism may facilitate word recognition of functional morphemes in allowing a broad inclusion of potentially quite phonetically aberrant forms. On the contrary, over-processing of functional morphemes could reveal novel acoustic variations that the listener fails to correctly match to previously heard tokens. The end result is that the ventral stream listening mechanism may allow the speaker to reduce functional class forms more freely, with these forms eventually becoming phonologized variants within the language.

One criticism of this argument would be that it relies on some degree of teleology, whereby speakers conciously shape the speech signal for some form of optimization. Ohala (1989) rejects such arguments, contending that innocent misperceptions on the part of the listener are the source of most sound changes. In this scenario, speakers may randomly produce a pool of variants, with no intention to reduce or hyperarticulate according to the needs of the listener. The listener, however, being aware of phenomena like coarticulation and reduction which may occur more
strongly in rapid, casual speech, is usually able to account for such reduction or coarticulation by perceptually undoing these effects and reconstructing the intended form of the word. In some cases, called hypocorrection, the listener fails to undo the effect fully, and perceives a new variant of the word. In some cases, the listener turned speaker produces the newly perceived form and a “mini sound change” has taken place.

Ohala’s model of listener-driven, non-teleological sound change may also fit in well with the dual stream model. Ventral stream listening may not necessarily only involve the lack of detailed phonetic awareness, but may also, or alternatively, involve more generalized listening that is prone to misperception, or less processing of phonetic detail, such that the effects of reduction and coarticulation are less likely to be fully undone (and may in turn be produced when the listener turns speaker). Thus when contextually less predicatable words are processed through the dorsal mode of listening, they may be less prone to misperception and non-faithful forms may be more accurately reconstructed by the listeners. Even a syntactic mapping of the two modes does not result in teleology; While it involves listeners learning an efficient means of processing speech for word recognition, neither the listener nor the speaker ever has any intention of optimizing the speech signal.

5.2. The Implementation Problem

Another potential problem with the dual stream model is that it predicts lower levels of phonetic awareness of more predictable word categories, which in turn allow for reduction of those variants; ultimately, how do those variants ever become part of the language if they are unnoticed? One possibility is that they are not entirely unnoticed, but simply that a larger number of tokens are required until speakers ultimately adopt them as new intended variants. Such a balance is required to maintain the slow pace of sound change that is generally observed. If functional
morphemes were both more likely to reduce and such reduction was readily noticed, adopted and retransmitted by listeners-turned-speakers, reduction would occur at an unbridled pace, as opposed to the slow erosion that tends to occur.

The second possibility stems from the proposal in 5.1. Ventral listening may also involve generalization rather than altogether lack of phonetic awareness. Ventral listening is said to involve the processing of words rather than segments, compared to dorsal listening, which shares a link to production and the segmentation needed for the articulatory mapping of sound (Hickok and Poeppel 2004). Generalization of phonetic characteristics might include broader, categorical perception of gradient allophones---such that a partially reduced [i] processed ventrally might be perceived either as [i] or [ə], but not something in between. A specific feature such as increased duration could be generalized as an increase in prominence, and in turn be perceived as an increase in intensity. The turbulence of a cough could be enough to qualify as a generalized [s] segment. Other types of misperceptions might be more prone to occurring in ventral listening, for example hearing nasalized vowels as having higher first formants or failing to undo articulatory effects. In whatever ways such a phenomenon may manifest itself, the model in the paper does not propose that lower phonetic attention necessarily means that such “ventrally perceived” tokens do not end up in the pool of tokens that may eventually affect a speaker’s production of those words.

5.3. Other Factors Affecting Dorsal and Ventral Processing

Without a doubt, other factors are involved in activating either the dorsal or ventral modes of listening besides syntactic positions associated with open and closed morphemes. For example, it is clear from Marslen-Wilson & Welsh (1978) that position within a word is important, and word beginnings likely merit higher degrees of dorsal listening than the ends of words. Likely the beginnings of sentence, or certainly conversations, would also involve dorsal stream listening,
where no semantic context yet accumulated. Cole & Jakamik’s (1980) observation about word recognition and density, where words with less common first syllables are recognized more quickly, presents a possibility that the listening mode is not fully determined by the syntactic structure, and that listeners may essentially deactivate dorsal stream listening once word recognition occurs. It is not clear just how active a listener would be in switching between modes, which could be quite a taxing process. Considering that ease of recognition could involve semantic context, syntactic predicability, position within a word or sentence, neighborhood density, etc., a perfect estimate of the listening requirement would involve quite a few calculations. Thus, the model proposed in the present paper does not rely on such calculations, and is instead a programmed mapping of syntax to listening mode. Future study can address whether, for example, phoneme restoration effects are stronger in the second syllables of words which are more easily identified from their first syllable (shampoo vs. conceive).

Another factor which could resolve the too-many-calculations issue is whether listeners can temporarily store words and ship them off to dorsal processing after determining that more precise listening is needed. This is potentially another problem with the proposed model, since “predictable” function morphemes often occur at the beginnings of phrases, and thus the beginnings of sentences. For example, in the sentence “The woman teaches linguistics,” in considering the entire sentence, the determiner ‘the’ would be highly predictable even without further context within a conversation. However, at the beginning of an utterance, with no context, the listener cannot make much of a predictable about the semantic context of syntactic structure in order to apply our model of listening. It is thus not clear if dorsal listening overrides the functional vs. content structure mapping utterance-initially, or if listeners begin with default ventral listening until some syntactic structure can be established, and are then able to “ship off” content words
through the dorsal processing route. Future research is needed to understand the interaction of these contextually divergent factors.

5.4. Dual Stream Processing and Noun/Verb Differences

A final topic of discussion concerns whether the dual stream processing model could be relevant in describing the differences in part of speech (nouns, verbs, etc.) phonologies. As shown in section 3, there is no evidence suggesting the activation of part of speech in the lexicon in such a way that would allow regular sound change to occur with category-specific environments. This does not rule out, however, that noun and verb phonological differences could be due to different listening modes at the syntactic level, which I have argued correlate with the reductive processes that occur with closed class morphemes.

Smith (2011) has argued for noun faithfulness, with more reductive processes occurring for verbs. Thus, it could possible that nouns warrant more precise phonetic listening through the dorsal stream, while verbs are processed more ventrally (although significantly less so than functional morphemes). Following the proposed analysis for the functional vs. content morpheme class distinction, the motivation for such a system would be in contrastive levels of contextual predictability. It is likely true that there are more nouns than verbs in the average language, although it is not clear how pronounced the difference would need to be in order to result in significantly different levels of predictability. Functional classes in English often have fewer than a hundred members, sometimes only a dozen or so as in the case of determiners, while content word classes have thousands, if not tens of thousands of members as in the case of nouns. Other characteristics of verbs---that they are more abstract (less likely to be imagable) and more difficult to learn in first or second language acquisition than nouns (Gentner 1981), suggests they have more in common with function words than nouns. Whether these characteristics would result in
different phonologies remains unanswered. While this seems unlikely since differences in contextual predictability between nouns and verbs is smaller, it is possible that the lower frequency of noun and verb specific phonological effects reflects this fact.

What factors then remain to account for the phonological differences that have been observed between nouns and verbs? While a detailed analysis is outside the scope of this paper, some studies have indicated that differences in phonological environments resulting from class-specific affixation as well as post-lexical phonological environment differences may be the cause of these differences. Blevins & Lynch (2009), for example, have shown how a sound change can appear to have class conditions when in fact there was a clear phonological environment historically. The environment is contained in affixes which are unique to verbal or nominal morphology, and once that environment later disappeared, the sound change can have the appearance of being class conditioned.

Other studies have examined cases of diffusive type changes which appear to occur at different rates for different categories. Brown (2013) investigates the case of /s/ reduction in New Mexican and Chihuahua Spanish which appears to be stronger in certain word classes. However, Brown measures the “frequency in a favorable context,” finding that word classes with higher occurrences of reduction correlated with a higher incidence of the phonologically favorable context. Essentially this is a post-lexical phonological environment, which considers the frequency of environments provided by neighboring words. Thus, these observations suggest that word class itself is not the influencing factor in the spread of the change, but rather usage-based factors such as “the phonetic environment in which different lexical categories tend to occur” (Bybee 2014:42). Further research is needed to investigate whether affix and post-lexical
phonological environment differences can account for the observed phonological asymmetries in nouns and verbs.

A final observation about noun-verb phonological differences is that they are primarily suprasegmental in nature rather than segmental. Furthermore, unlike the differences seen between functional and content morphemes, noun-verb asymmetries rarely show the same type of reduction. Although I argue that content and functional morpheme phonological asymmetries originate in prosodic differences—-a decrease in phonetic prominence, primarily—-such prosodic differences tend to result in the reduction of segmental contrasts, consonant lenition, or vowel centralization. While Smith (2011) shows evidence of a bias for contrast reduction in verbs, it is typically in the number of tonal or pitch accent contrasts or of other suprasegmental features. However, there are some counterexamples of verb privilege as well, while reduction targeting content but not functional morphemes would be exceedingly rare. This evidence suggests nouns and verbs are merely interacting with the syntax based on different distributions. This may also provide support for a hypothesis that Phillips’ (2006) examples of class-specific lexical diffusion asymmetries might be due to Brown’s (2013) and Bybee’s (2014) frequency of favorable condition, if nouns and verbs only exhibit differences based on syntactic level phonological environments.

6. Conclusion

Category-specific phonology is a well-documented phenomenon, occurring commonly in function vs. content morpheme systems but is less frequent in distinguishing more specific categories like nouns and verbs. Studies in neuro- and psycholinguistics have revealed profound
differences in the way in which nouns and verbs, as well as open and closed class morphemes, are processed. This paper has reviewed the literature in order to determine if there is a link between category-specific cognitive differences and category-specific phonology. Ultimately, I have shown that there is little to no evidence to suggest categories are active within the lexicon, which I consider a key criteria for either of two hypotheses: that there are innate or universal constraints guiding these changes, or that words of shared categories promote diffusion of sound changes within those categories.

Despite this, I have argued that different listening types, dorsal ‘how’ listening vs. ventral ‘what’ listening, are mapped to content and functional categories at the syntactic level. This proposal is deduced from the strong correlation between low predictability, a higher level of phonetic attention, and increased faithfulness, as opposed to higher predictability, a lower level of phonetic attention, and the tendency for reduction. Differences in noun and verb phonology may also arise at the syntactic level, but are likely due to usage-based factors such as category-specific affixation and tendencies of post-lexical phonological environments. Additional details, such as the role of the speaker in relation to the listener, the interaction of multiple environments warranting different listening modes, and the issue of implementation, have as of yet only speculative accounts, but future research into relationship between listening modes and sound change will help answer these questions.

7. Bibliography


Smith, Jennifer L.  2012.  ‘Parts of speech in phonology,’  Handout from presentation at the 17th LIPP Symposium “Parts of Speech across Languages, in Acquisition, Mind and Brain”;” LMU, Munich, July 6.


