Title

Combining Structure and Usage Patterns in Morpheme Production: Probabilistic Effects of Sentence Context and Inflectional Paradigms

Author
Cohen, Clara

Publication Date
2014

Combining structure and usage patterns in morpheme production: Probabilistic effects of sentence context and inflectional paradigms

by

Clara Philena Cohen

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Linguistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Susanne Gahl, Chair
Professor Keith Johnson
Professor Line Mikkelsen
Professor Darya Kavitskaya

Spring 2014

Combining structure and usage patterns in morpheme production: Probabilistic effects of sentence context and inflectional paradigms

Abstract

Combining structure and usage patterns in morpheme production: Probabilistic effects of sentence context and inflectional paradigms

by

Clara Philena Cohen

Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Susanne Gahl, Chair

In this dissertation, I ask how systematic patterns of pronunciation variation in speech production reveal speakers' awareness of abstract structure and usage patterns during the planning and articulation of an utterance. I explore how structure and usage combine to affect speech production at the level of the sentence and the level of the word. In particular, I focus on subject verb agreement, because computing agreement draws on sentence-level mechanisms and word-level mechanisms of speech production. At the sentence level, agreement implicates structural knowledge because speakers must know which phrases are subjects and which words are verbs in order to inflect the verb appropriately. It implicates usage patterns because, in some constructions, multiple agreement forms may be used, each with its own probability of being selected. The contextual probability of using a particular form in such constructions is thus the sentence-level intersection of structure and usage. At the word level, agreement implicates structural knowledge through paradigmatic relations between inflectional forms of a lexeme. Speakers must know that the relationship between *speak* and *spoke*, for example, is the same as the relationship between *talk* and *talked*, despite the lack of any phonological similarity between the two alternations. Usage patterns come into play at the word level because different forms within a given lexeme's paradigm are used with vastly different frequencies, and the shape of these frequency distributions affects how people retrieve words (Bien et al., 2011; Kuperman et al., 2007). The paradigmatic probability of using a particular form from an inflectional paradigm therefore represents the word-level intersection of structure and usage. In three experiments in English and Russian, I ask how contextual and paradigmatic probability of using a particular form affect the articulation of that form.

In both English and Russian, I find that higher contextual probability of producing a particular agreeing form results in some type of phonetic reduction, while higher paradigmatic probability results in some type of phonetic enhancement. The phonetic feature that shows reduction or enhancement, however, depends on the language. This feature specificity leads me to propose the Contrast Dependent Pronunciation Variation hypothesis (CDPV). According to this hypothesis, structure and usage combine to restrict the types of probabilistic pronunciation variation that a speaker employs. By CDPV, pronunciation is not simply "reduced' when certain forms are con-

textually probable (Jurafsky et al., 2001) or "enhanced" when certain forms are paradigmatically probable (Kuperman et al., 2007). Rather, the phonetic features that vary with respect to contextual or paradigmatic probability are exactly those features which encode salient contrasts between competing forms. Phonetic "reduction" and "enhancement" are not general processes that weaken or strengthen the articulation in predictable, universal ways. Rather, they are targeted adjustments — reductions and enhancement — of the contrasts themselves, and are therefore sensitive to language-specific and perhaps even construction-specific properties.

To Daniel
(who else?)

Contents

List of Figures

List of Tables

Acknowledgments

One of my favorite things about graduate school is how freely everyone will share their expertise for the asking. This dissertation could not have been written without the help of Johanna Nichols and Nina Dobrushina, who were key in my struggle to find native Russian speakers living in Russia. My research assistants Daria Barylnikova and Bella Yashayev deserve special mention: Dasha for her on-site aid in the logistics of running the experiment in Chapter 3, and Bella for her assistance in stimulus design and data analysis for the work in Chapter 4. I am also grateful to everyone at the Higher School of Economics in Moscow — students, faculty, family members of the above, and everyone else I met. They were kind and helpful and patient as I recruited experimental participants and searched for quiet rooms to run the experiment, and I will remember their hospitality long after I forget the specifics of this dissertation.

I also cannot forget the many native Russian speakers in the SF Bay area who patiently helped me improve my extremely laughable attempts to construct experimental stimuli. Oleg Mayba, after reading my initial attempts, compared them to "Soviet realist poetry" (much kinder than "laughably bad") before rewriting most of them entirely. Further improvements came from Svetlana Taubinsky, Olga Dmitrieva (who also willingly translated my consent forms for me), and my outside committee member Dasha Kavitskaya. Despite their efforts, correcting my sentences was like herding cats, and experimental constraints, combined with last-minute adjustments, meant that many of these stimuli were odd or downright ungrammatical. All stray cats are my responsibility alone.

The aid I received stretched beyond all things Russian. Natalie Babler, Jane Paris, Belén Flores, Paula Floro, and Ron Sprouse could fix *anything*, be it a class room change, a broken computer, a problem with fellowships or filing deadlines, or my unpleasantly frequent habit of locking myself out of my office. Both Victor Kuperman and Harald Baayen were very gracious in lending extensive help in refining my statistical analyses, and I cannot forget Jordana Heller, who was my partner at the LSA Summer Institute at Boulder, where I first developed this project.

Of special value were my academic peers — both my office mates on the C-level and my cohort-mates on the A-level. Melinda Fricke showed me that it can be done; John Sylak-Glassman and Stephanie Farmer were always willing to chat whenever I wandered by the Language Documentation Office; Andrea Davis and Shinae Kang would always lend an ear whenever I needed to talk through an analysis; Sarah Bakst was a constant source of enthusiasm and good humor at the neighboring desk; and Emily Cibelli somehow made me feel more sane by radiating serenity and confidence whenever I was in her presence. And Stephanie Shih was an inspiration in everything she did, from her quantitative abilities to her mastery of all things baked. I will always value our ladies' teas with everyone associated with the C-level.

My committee was everything a committee should be. Line, thank you for your constant encouragement, your speedy and complete feedback, and the model you provided of the sort of scholar and teacher I would like to be. I almost became a formal syntactician because of you, and my interest in agreement is largely the result of your influence. Keith, thank you for your relaxed good humor, which served as a frequent reminder that there's really no need to be so stressed out. Dasha, I have already thanked you for your help in constructing my stimuli, but you deserve a second mention. You were my checkpoint for all things Russian, and I am very grateful to you for catching all my errors — both grammatical and (more embarrassingly) phonological. And, Susanne, I decided to become a psycholinguist after taking Ling 160 with you. Thank you for guiding me through the process of becoming a psycholinguist, for teaching me how to analyze data, for telling me to learn Perl, for making me suck it up and rewrite everything that needed rewriting, and for all of your support and encouragement (and networking on my behalf) as I started reaching out beyond the safe walls of Dwinelle and engaging with everyone else in my field.

Daniel, I saved you for the last two lines, because, as Billy Collins tells us, that's where it really counts.

Chapter 1

Introduction: Structure and usage in language

1.1 Introduction

A common example in introductory linguistics classes is Lewis Carroll's poem *Jabberwocky*, a composition of meaningless gibberish that clearly illustrates the disconnect between two different types of patterns that characterize language. Consider the first verse:

(1) 'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves
And the mome raths outgrabe. (Carroll, 1871)

There are no recognizable content words in these lines, making it impossible to extract concrete meaning. Yet the poem is recognizably English, identifiable based on the combination of function words, morphological inflections, and word order. In the third line, for example, *All mimsy were the borogoves*, the use of the plural form *were* indicates that the subject of the sentence is probably plural. The following words, *the borogoves*, contain an initial determiner, *the*, indicating a noun phrase, and the final *-s* on *borogoves* is sufficiently similar to a plural suffix to yield the conclusion that this phrase serves as the subject of the sentence. Regardless of novelty of the vocabulary and the non-standard post-verbal position of *the borogoves*, it is still possible to decode the underlying structure of the sentences and words because of our knowledge of the structural patterns that underlie language. Phenomena such as the combinatorial patterns of determiners and morphosyntactic indicators of subject-verb agreement are key cues to these structural underpinnings.

Jabberwocky, with its grammatical combinations of novel words, is one demonstration of the independent existence of a structural skeleton. Another is Chomsky (1957)'s famous example *Colorless green ideas sleep furiously*, which illustrates a grammatical combination of existing words that nevertheless fail to provide semantic content. These examples are recognizably English not because of their use of English vocabulary (largely absent in *Jabberwocky*) or because of their successful conveyance of an intended message (entirely absent from *colorless green ideas* ...). They are recognizably English because they employ English grammatical structures.

Yet a language is characterized by more than the set of possible grammatical structures. Despite the ability to create unique utterances demonstrated so memorably in *Jabberwocky*, people rarely exploit these combinatorial resources to their full capacity. A key part of knowing language is knowing which utterances tend to be used more frequently than others, a pattern that emerges not as a result of the structural capacity of a language, but because people tend to reuse useful utterances. Consider, for example, the plural *were* in line 3 of (1). The structural underpinnings of English license this verb form in two cases: either an indicative sentence with a second person or plural subject (*We/you/they were*), or in particular irrealis constructions, regardless of subject (*He wishes he were dead/I wish I were dead*). Yet of the two constructions, the indicative clause with a plural subject is far, far more common than the irrealis. It is this knowledge that encourages the interpretation of *were* as a plural indicative verb. Despite the fact that English's structural capacity licenses the word form in two different ways, usage patterns tell us that one is much, much more common than the other.

The interaction of these two types of linguistic patterns — structural patterns and usage patterns — is the focus of this dissertation. This type of distinction is hardly new: From Saussure's distinction between *langue* and *parole* (de Saussure, 1959) to Noam Chomsky's contrast between competence and performance (Chomsky, 1965), linguists have long been aware of the difference between linguistic patterns that can be described with a certain degree of abstract generalization, and patterns of how people actually use language. These early views, however, made the distinction between abstract structural properties and repeating usage patterns not as two patterns that contribute to human language, but as one pattern that defined language, *langue* or competence, and a second layer, *parole* or performance, that overlaid language. This second layer was not necessarily structured or patterned in any systematic way; it served only to obscure the pure linguistic structure that defined language. These early approaches addressed usage properties only because it was necessary to know what to discard in order to reveal the abstract structural properties of language, which formed the actual topic of study.

More recent work, however, has pointed out that usage patterns are, indeed, systematic, structured, and can be informative about human language. Cedergren and Sankoff (1974), for example, observed that certain processes, such as *r*-spirantization in Panamanian Spanish, do not apply categorically, but probabilistically; they proposed that these probabilities of occurrence should be added as a component to linguistic competence. Later theories went farther, treating abstract structure not as the pure substance of language, but as a secondary, emergent property that children must learn by generalizing from observed usage patterns. By these theories, language acquisition is not a question of determining the correct setting of a universal set of syntactic parameters (e.g., Baker, 2008) or the correct ordering of a universal set of phonological constraints (Prince and Smolensky, 2004). Rather, language learners all acquire the abstract structural patterns that emerge in natural language by inferring them in repeated usage patterns that they observe with no pre-existing linguistic framework (Barlow and Kemmer, 2000; Bybee, 2007). By these views, the adult grammar is a combination of structural patterns, which have been arrived at through abstracting over usage patterns.

The view that I will adopt is closest to that of Cedergren and Sankoff (1974), or the views by which structural patterns are emergent generalizations (Barlow and Kemmer, 2000; Bybee, 2007). In other words, I will assume that people have knowledge of both the abstract combinatorial possibilities of language (structure) and of the rates at which certain combinations (or types of combinations) are actually employed (usage). This dichotomy is similar to the divisions between *langue* and *parole*, or between competence and performance, because I assume a qualitative distinction between these two types of linguistic properties, such that structure overlaps largely with *langue* or competence, while usage overlaps largely with *parole* or performance. The key difference between those divisions and my own is that a person's knowledge of a language includes both of these properties. I do not propose to study the effects of structure alone, and discard usage as epiphenomenal. To understand how people produce language, it is necessary to understand how they integrate both types of knowledge, and that is the goal of this dissertation.

To explore the integration of structure and usage, it is first necessary to have a clear understanding of the assumptions that I make about them. In the remainder of this section, I will specify the properties of structure and usage patterns, as I see them, that are of key interest in this dissertation. In particular, I focus on two domains of interaction — sentence-level and the word-level — and I lay out how subject-verb agreement suffixes are the locus of structure and usage combinations at both levels. In Section 2, I review some theories of production of agreement at the sentence level, and in Section 3 I summarize two theories of how morphologically complex words are retrieved. In Section 4 I lay out how these theories can yield quantitative measures of the combination of structure and usage, which will form key variables of interest in the experimental portion of this dissertation. In Section 5 I describe pronunciation variation, the key investigative tool that I employ as a proxy measure of how words and sentences are produced. Finally in Section 6 I lay out the three production studies that comprise the core of this dissertation, in which I examine the production of stubject-verb agreement in English and Russian.

1.1.1 What is structure?

Language is flexible. By making use of a finite set of base units — phonemes, words, morphemes — and a finite set of patterns by which we combine them, we are capable of expressing an infinite range of meanings. These combinatorial patterns make up a language's grammar, and one primary goal of linguistic research is to understand and describe as precisely as possible the exact nature of a language's grammar. It is this research that has yielded modern understandings of linguistic structure ranging from phonotactic constraints (combinatorial patterns of individual sounds), to morphological structure (combinatorial patterns of morphemes), to syntax (combinatorial patterns of words), and even logical semantics (combinatorial patterns of denotation). Without these combinatorial patterns, compositionality — the construction of novel phrases and the expression of novel ideas — would be impossible. These patterns, which I will combine under the general heading of *structure*, are absolutely necessary to language.

Structure is present in multiple domains. One classic example is syntactic ambiguity, in which multiple meanings of a sentence correspond to multiple syntactic structures. For example, the sentence *One morning I shot an elephant in my pyjamas* (Heerman, 1930) can have two distinct syntactic structures. In one case, the prepositional phrase *in my pyjamas* is associated with the noun phrase headed by *elephant*, yielding the interpretation that the elephant is in the speaker's pyjamas. In the other case, the prepositional phrase is associated with the verb phrase headed by *shot*, yielding the interpretation that the speaker is in his pyjamas, and the elephant is unclothed. Two distinct meanings are expressed using the same words in the same order, simply by making use of two different structures.

Subject-verb agreement is also an instance of structural patterns that are active at the sentence level. The distinction between verb forms such as *is* and *are* in English is determined almost entirely by the fact that there is a structural link between the linguistic entity that plays the role of the subject of the sentence and the linguistic entity that plays the role of the verb. Regardless of the intended referent or the lexical item used to express these entities, the connection exists.

Structural patterns also help organize complex word-formation. English inflectional morphology, for example, can be described as consisting of, minimally, a root, on which inflectional morphology regularly occurs word-finally. As long as the root is regular, it makes no difference which morphemes are involved. Any root and any inflectional morpheme will combine in this order, because that is the structural pattern that governs inflection. Moving up, from the level of the word form to the level of the lexeme — all inflectionally related word forms — we can see another sort of structural pattern: The inflectional paradigm. This paradigm is the set of word forms that a given lexeme can take, and a key property of an inflectional paradigm is that all forms are available for all lexemes. In this way, verb pairs like *speak/spoke* and *walk/walked* are understood to index the same type of alternation — namely, present and past tense — regardless of the lack of formal similarity between the two alternation patterns, because the two pairs occupy the same two cells in the inflectional paradigm. This systematicity in possible forms is another structural pattern that governs inflection.

This dissertation will be concerned mainly with sentence-level and word-level structural patterns, but it should be noted that structural patterns pervade language at all levels. At the sub-lexical level, it is possible to analyze individual sounds as bundles of distinctive features, such as [+voice] or [LABIAL], which combine in different ways to create the phonemes of a given language, phonemes which themselves combine according to language-specific patterns to create morphemes (Chomsky and Halle, 1968; Clements, 1985). At the discourse level, different intonational contours can be used to convey specific abstract discourse functions, such as polar questions or contrastive focus, regardless of the actual syntactic structure or semantic content of the utterance (Pierrehumbert and Hirschberg, 1990). Combined, all of these structural patterns contribute to the grammar of a language. It is because of them that sentences such as *All mimsy were the borogoves* are recognizably English. Despite the fact that two of the key words are not part of the English lexicon, they are nevertheless made up of English phonemes, combined according to phonotactic patterns, decomposable into English morphemes, and ordered according to English syntax. The structural patterns of English are responsible for the fact that this sentence could be considered grammatical, despite the fact that no one ever would use it.

1.1.2 What is usage?

In contrast to structure patterns, usage patterns are emergent from the sea of utterances that all people are exposed to and contribute to throughout their entire lives. Whereas structural patterns define abstract frameworks that speakers use to combine whichever linguistic units they wish to employ, usage patterns are defined by the rates at which speakers actually do employ specific combinations of these units.

As with structural patterns, usage patterns pervade all domains of language, from the frequencies with which certain phonemes are used in various positions within the word (Vitevitch and Luce, 2004) to the rates at which people actual produce questions with question intonation (only about half the time; Geluykens, 1988). At the sentence level, between 20% and 25% of multiword utterances are recognizably recycled, formulaic expressions (Sprenger, 2003; Van Lancker-Sidtis and Rallon, 2004), and even when those expressions are not identifiable formulas, it is still the case that different combinations of words occur together far more or less often than would be predicted based on their baseline frequencies (e.g., Jurafsky et al., 2001). These baseline frequencies themselves represent their own usage patterns at the word level. By taking into account the frequencies of individual words, researchers have uncovered relationships between word frequency and the structure of a language's lexicon (Zipf, 1935), a word's phonological realization (Fidelholtz, 1975), or the productivity of a particular morphological pattern (Baayen, 2009; Bybee and Slobin, 1982; Hay and Baayen, 2001).

The fact that these usage patterns emerge so markedly from language data has led to an alternative approach to linguistic research, in which these usage patterns are the driving force behind models of language (e.g., Barlow and Kemmer, 2000; Bybee, 2007). Rather than describing language in terms of a lexicon and syntactic structures, or a set of morphemes and patterns by which they combine to form words, these models describe apparent abstract structures as emerging patterns that can be generalized from repeatedly used construction. Usage patterns are how we know that, regardless of whether *all mimsy were the borogoves* is consistent with structural patterns of English sentences and words, no one would ever produce it. And usage patterns are how we know that, even if *mimsy* and *borogoves* were real lexical items, this type of inversion of subject and predicate is uncommon and used mainly in high-style, poetic registers.

1.1.3 Combining structure and usage in verb agreement

The goal of this dissertation is to explore how structure and usage are integrated by the speaker during speech production, both at the level of the sentence and the level of the word. In order to do that, it is necessary to identify some phenomena that are sensitive to both types of patterns, at both levels. Subject verb agreement provides just such a case.

At the sentence-level, psycholinguistic research has demonstrated that language users are sensitive to both structure and usage when they compute subject verb agreement. One phenomenon that illustrates how these types of information can be combined is agreement attraction, a phenomenon which will be of key interest in this dissertation. Agreement attraction refers to a particular type of speech error in which a predicate agrees with some noun that is not the head noun of the subject noun phrase. For example, in the sentence *The science of speech waves were well understood*, uttered during an informal presentation on the history of phonetics (31 March 2014), the verb *were* shows plural agreement, presumably triggered erroneously by the plural *speech waves*, rather than singular agreement as would be required by the singular head of the subject, *science*. The rate at which people produce plural agreement in these constructions is sensitive to many properties of the sentence and discourse context, and those properties can both be structural or the result of a usage pattern. One structural property that affects the rate of agreement attraction is the syntactic position of the attracting noun. Franck et al. (2002), for example, found that agreement attraction occurs more frequently when the non-subject noun is less deeply embedded in the subject noun phrase, and Vigliocco and Nicol (1998) showed that this effect is independent of the linear proximity between the verb and the non-subject noun. Haskell et al. (2010) observed that usage patterns of the discourse context can also have an effect, by showing that agreement attraction can be primed. People will produce more erroneously plural verbs when they have seen more plural verbs used with similar subjects. For these reasons, agreement attraction is an excellent testing ground for exploring how structural and usage patterns interact, because the rate at which it occurs is sensitive both to the speaker's knowledge of the structural pattern of the utterance, and to the knowledge of the usage patterns of those agreement forms.

At the word-level, agreement is of interest because it is implemented through morphologically complex words in English and Russian, the languages studied in this dissertation. Like sentences with agreement attraction constructions, morphologically complex words also show sensitivity to structure and usage. Masked priming experiments, for example, have shown that people automatically decompose words into possible constituent morphemes, regardless of whether those morphemes are actually present in the word (e.g., *corner* → *corn* + *er*) (Kazanina, 2011; Kazanina et al., 2008; Rastle et al., 2004). This indicates that people draw on their knowledge of possible morphological structure, at least in the earliest stages of lexical retrieval, regardless of whether those structures are actually used in the word. Yet they are also acutely aware of the usage patterns associated with morphologically related words. When people are asked to rank pairs of morphologically complex English words as more and less complex, their decisions are affected by the relative usage frequency of the complex word and its base morpheme (Hay, 2001). When people were asked to produce complex Dutch adjectives, they were faster when the base morpheme of the adjective was more frequent overall (Bien et al., 2011), and when they were asked to describe pictures using Dutch verbs, their responses showed an awareness of inflectional entropy, or how similar the usage frequencies of all the inflectionally related forms were to each other (Tabak et al., 2010). Together, these findings show that people are acutely aware both of possible and actual morphological structures, and also of the usage patterns associated with the individual morphological components of complex words.

Variation in the selection of an agreeing verb form represents the combination of both structure and usage. At the sentence level, this variation is a subset of the class of phenomena known as *grammatical variation*, wherein the selection of a particular structural pattern can be described according to a given usage pattern. In English alone, such cases include usage patterns such as how often a verb particle occurs immediately after the verb (Gries, 2003), how often people realize the dative alternation with a double object construction instead of a noun phrase followed by a prepositional phrase (Bresnan et al., 2007), how often people use the *s*-genitive rather than the *of*-genitive (Szmrecsanyi and Hinrichs, 2008), how often people use active voice rather than passive voice (Weiner and Labov, 1983), how often people use *which* as a relative pronoun, rather than *that* (Guy and Bayley, 1995), and how often existential constructions exhibit number agreement (Hay and Schreier, 2004; Riordan, 2007). For any given type of grammatical variation, the probability of observing one variant over another represents the combination of structure and usage, because it is derived from the usage patterns of using particular structures. In the case of agreement, the structure of interest is a particular agreement relation between subject and verb, which is indexed by the inflectional form — such as third person singular — that appears on the verb. The usage pattern of interest is the rate at which people produce that particular form of the verb.

At the word level, combinations of structure and usage appear in the form of relative frequencies within an inflectional paradigm. As described above, inflectional paradigms are another type of pure structure that relates different forms of a lexeme to each other. They can be described as sets of empty cells that are filled by the various forms of a lexeme of interest, independently of the identity of the lexeme or the frequencies with which those forms are used. Yet the rates at which people might need to use the occupant of a particular paradigmatic cell do vary depending on the verb in question. As a result, the relative frequency within an inflectional paradigm of a particular form of a verb — such as third person singular — is another combination of structure and usage. Subject-verb agreement variation therefore is the phenomenon that permits investigation of how people combine structure and usage patterns both at the sentence and word level.

1.2 Theories of agreement production at the sentence level

In this dissertation, I examine the production of agreement morphology in English and Russian sentences that show variation between singular and plural agreement. The construction in the two languages that permits this variation, however, is not the same. In English, I use agreement attraction, while in Russian, I exploit quantified subject noun phrases. This section serves to lay out the theories of production that have been developed to account for the variability occuring with both types of construction.

1.2.1 Agreement attraction

Agreement attraction is the phenomenon whereby predicates agree with nouns that are not the head noun of the subject. Such nouns, often called *local nouns* because they are usually linearly closer to the verb than the intended controller, "attract" agreement away from the subject. The classic example of agreement attraction involves number, shown below in (2).

(2) The key to the cabinets is/are rusty (Bock and Miller, 1991)

In this type of sentence, the subject's head noun *key* is singular, yet the plural local noun *cabinets*, induces plural agreement on the verb (*are*) about 13% of the time on average (Eberhard et al., 2005). This phenomenon has been heavily studied, yielding a number of theories regarding how people resolve the link between subjects and predicates. Most frequently it is described as a type of speech error — as, indeed, it is considered to be in standard English — but according to the theory that will be adopted here (Eberhard et al., 2005), the mechanism that produces it is the same type of probabilistic grammatical variation that can be observed in the dative alternation. Further, by this theory, there is no difference between the production process involved in producing an "erroneous" plural verb and a "grammatical" singular verb, a claim that has been supported by response time evidence (Staub, 2009, 2010).

Some of the earliest work on agreement attraction has focused on the importance of the syntactic position of the local noun with respect to the verb. Bock and Cutting (1992) found that plural nouns embedded in a subordinate clause (e.g., *the actor [who directed the **films**]*) elicit much lower rates of plural verbs than those same nouns embedded in PPs (e.g., *the actor [in the blockbuster **films**]*). They propose that this effect is due to the fact that the core planning unit of grammatical processing is the clause. Local nouns in subordinate clauses are not in the same clause that contains the head noun and the verb; this insulation from the main clause limits their ability to interfere in grammatical planning processes in that clause, such as agreement. Yet this clausal insulation was not complete: Bock and Cutting (1992) did still find attraction even with local nouns embedded in clauses. Later work proposed that the important distinction is not a binary issue of clause-packaging, but rather a more gradient case of syntactic distance. Vigliocco and Nicol (1998) found that there was no difference in agreement attraction rates between sentences such as *The helicopter for the flights is/are safe* and their corresponding polar questions (*Is/Are the helicopter for the flights safe?*). They proposed that agreement attraction is the result of syntactic proximity between the local noun and the verb. This proposal was supported by the findings of Franck et al. (2002), who observed that a plural attractor in an intermediately embedded position (e.g., *The helicopter [for the **flights** [over the canyon]]*) elicited more plural agreement than local nouns in a deeply embedded position (*The helicopter [for the flight [over the **canyons**]]*), even though the latter were immediately adjacent to the verb. Hartsuiker et al. (2001) further refined this proposal to show that the key syntactic distance is the separation between the local noun and the head node of the subject noun phrase, rather than between the local noun and the verb. All of this work supports a model of agreement in which the transfer of number features from a controlling noun to a verb is mediated by the syntactic structure of the sentence. The closer the attractor noun is to the head node of the noun phrase, the more easily the spurious plural features can percolate up and accidentally control the verb agreement.

Other work on agreement attraction, however, has suggested that syntactic structure is not the only property which affects how agreement is computed. Thornton and MacDonald (2003), for example, found that people make significantly more attraction errors if the local noun is a plausible subject for the verb. For example *The album by the classical composers was praised* elicited more plural agreement than when the verb was *played*, as composers are not typically played, but are frequently praised. Haskell and MacDonald (2003) found that the morphophonologically regular local nouns (e.g., *rats* but not *mice*) elicited more plural agreement only when semantic factors and grammatical factors were in conflict. Thus, *rats* induced more plural agreement than *mice* only when it was the local noun for grammatically singular but notionally plural head noun, as in *the family of rats*. There was no such effect when the notional and grammatical number of the head noun matched, as in *the cage for the spotted mice/rats*. Finally Haskell et al. (2010) showed that agreement attraction can be primed. MacDonald and colleagues account for these findings with a constraint satisfaction model of agreement, in which the inflected form of the verb is selected according to a number of different cues, which have different strengths. One such cue is previous experience with subject-verb associations: When speakers have previously encountered certain nouns as subjects for certain verbs, this experience provides a strong cue that those nouns should once more be subjects. For this reason, plausible local nouns elicit more agreement attraction than implausible ones. A second property is that the various strengths of cues can interact, so that competition between contradictory strong cues, such as grammatical and notional number, can make it possible for weaker cues, such as morphophonological regularity, to show an effect. This is why local *rats* elicited more plural agreement than local *mice* only when the head noun was a collective. MacDonald and colleagues propose that this model can account for the existence of agreement attraction in the first place. Like Staub (2009, 2010), they do not consider agreement attraction a speech error, in which something has gone wrong during the planning or utterance of the sentence. In their account, agreement attraction is a general consequence from previous experience with collective expressions such as *a number of considerations*, *a series of consequences*, which are headed by a singular noun, yet freely occur with plural agreement. It is this usage pattern, they suggest, that provides a cue associating subjects of the form $[N_{sg}\ P\ N_{pl}]$ with plural agreement.

This cue-based approach also appears in work by Franck et al. (2008) investigating gender agreement attraction in French, Italian, and Spanish. Thornton and colleagues' genesis account of number agreement attraction based on experience with "number of" expressions cannot account for gender agreement attraction phenomena in other languages, but their constraint satisfaction model with weighted cues does transfer. Franck et al. (2008) found that the form of the head noun determiner had a stronger effect on agreement error rates in French than in Spanish, and no effect at all in Italian. They attribute this difference to the fact that the endings on the nouns themselves provide strong cues to head noun gender in Italian, less in Spanish, but very little at all in French. This means that, in Italian, it is possible to retrieve the subject's gender based entirely on phonological information in the noun. In French, by contrast, there is very little systematic relationship between noun endings and gender information, so speakers must draw upon the form of the determiner as a cue. In other words, speakers are sensitive to the validity of gender cues in their language, and rely on those cues to signal the required predicate agreement form in proportion to the information that their previous experience has provided them.

These findings demonstrate that both structure and usage affect agreement. There are a number of models laying out how people compute number agreement during production, but to my knowledge only two take into account both structural and usage information. The first, the Maximal Input hypothesis, was proposed by Vigliocco and Hartsuiker (2002). According to this model, all information is available at every stage of language production, but it is weighted according to its reliability. It is this sort of model that explains how morphophonological information about noun gender in French, Spanish, and Italian can affect predicate agreement (Franck et al., 2008), or how morphophonological regularity can have (very) small effects on number agreement in English (Haskell and MacDonald, 2003): Information about the form of the noun is available during the computation of the predicate's form, and the extent to which that information is used is a function of how experience with that information has revealed its reliability.

A second model that incorporates both structure and usage is Marking and Morphing (Bock et al., 2001; Eberhard et al., 2005), In this model, unlike Maximal Input, information which is relevant at one stage of production does not affect the second stage. This model distinguishes two primary stages of agreement computation. In the first stage, *marking*, subject noun phrases are marked for notional number, information which comes from the conceptual message that the speaker is preparing for articulation. It is at this stage that conceptual number has the opportunity to affect agreement (Humphreys and Bock, 2005; Vigliocco et al., 1996b). The second stage, *morphing*, is the point at which individual words' morphological constituents are specified. This is the stage at which plural morphology on local nouns can become active and interfere with the notional marking on the head noun. The output of morphing is a number specification on the subject noun phrase that is the combination of the notional number assigned during marking and the morphological number assigned during morphing. It is this number specification that is transmitted to the verb and determines the agreement morphology.

According to this model, agreement attraction effects arise at the second stage of grammatical planning, and are the result of "the dynamics of morphing run amok" (Eberhard et al., 2005, pg. 10). For agreement to occur, the number features of the entire subject noun phrase must be specified and transmitted to the verb during morphing. When a noun phrase that has been marked as singular contains another element that is inflected as plural (such as a local noun), the plural morphology on the subordinate element can affect the eventual number specification of the entire subject. If the plural attractor is strong enough to override the head noun's number features, then the information transmitted to the verb will reflect the plurality of the attractor, rather than the singularity of the head noun. A key property is that the ability of the attractor to override the singular head noun is not binary. The model assigns to subject noun phrases a gradient plurality specification, which corresponds to a probability of producing plural agreement. Staub (2009, 2010) provides reaction time evidence showing that when people encounter agreement attraction constructions, they are slower to produce a verb than when the choice is unambiguous and, crucially, they are slowed by the same amount regardless of whether they produce a "correct" singular verb, or an "erroneous" plural. This suggests that the processes by which singular and plural verbs are produced are not qualitatively different. It is not the case that one is an error, and produced differently from the other; rather, both singular and plural agreement are produced by the same grammatical planning mechanism, and the choice between them is probabilistic.

This model incorporates both structural and usage information to constrain how easily the plural attractor can override the head noun's number features. There are two primary factors that can have an effect: Syntactic distance (structure) and contrastive frequency (usage). Syntactic structure constrains agreement attraction rates because an attractor's number specifications cannot propagate as easily over a large syntactic distance. This is the source of the syntactic distance effects found by Bock and Cutting (1992); Franck et al. (2002); Vigliocco and Nicol (1998) and Hartsuiker et al. (2001). Usage patterns are incorporated into the model to account for the fact that the frequency with which an attractor noun is used in singular or plural forms affects the rate of ageement attraction. Local nouns which are used in both singular and plural forms, like *bubble/bubbles*, are stronger attractors than invariant plural nouns which have no singular counterparts, like *scissors* or *suds* (Bock et al., 2001).

The model is by no means ideal. It does not explain how agreement attraction can be primed (Haskell et al., 2010), and it cannot explain the effect of plausibility (Thornton and MacDonald, 2003). It cannot account for the effect of morphophonology on agreement attraction, as has been found in Romance languages (Franck et al., 2008) and Serbian (Mirković and MacDonald, 2013), because it has no mechanism by which low-level morphophonological information can affect higher level morphosyntactic relations. Its key contribution, however, is the fact that it has an explicit computational implementation that has been rigorously normed on English data. For research on English agreement attraction, therefore, it is superior to all other models, because it allows for the computation of a precise probability of observing plural agreement for any sentence, which takes into account the syntactic structure of the sentence, as well as the usage patterns of the individual nouns that make up the subject. It calculates predicted usage patterns of the verb forms, and therefore makes it possible to analyze how production varies as a given usage pattern varies. For this reason, it will be the basis of the analysis of sentence-level agreement production in English, presented in Chapter 2.

1.2.2 Quantified subject noun phrases in Russian

In Russian, sentences with quantified subjects (e.g., *four books, many boys*) can take either singular or plural agreement, depending on a large variety of factors, including the quantifier involved, the animacy of the subject, the information structure of the utterance, and the tense and semantics of the predicate (Corbett, 1979, 1983, 2006; Kuvšinskaja, 2012; Lambrecht and Polinsky, 1997; Nichols et al., 1980; Patton, 1969; Robblee, 1993a, 1997; Suprun, 1969; Timberlake, 2004).

Accounts of this particular form of agreement variation tend to look for structural differences between sentences preferring plural agreement and those preferring singular agreement. One such approach can be found in Corbett (1983)'s account of the effect of quantifier on agreement patterns — specifically, the fact that larger numerals tend to condition singular agreement. Thus, paucal numerals — i.e., *dva/dve, tri, četyre* (two, three and four) — are far more likely to show plural agreement, while higher numbers, such as those between 5-10, can vary between singular and plural more freely, and extremely large numbers, such as *tysjača*, 'one thousand,' must take singular agreement.

Corbett explains this pattern by appealing to the structure of the lexical categories of the quantifiers. He observes that the set of numerals from one through one million vary in other syntactic properties, resulting in extremely adjectival behavior on the side of lower numbers, and extremely nominal behavior in the side of large numbers. Table 1.1 repeats his matrix of agreement properties, some of which are characteristic of adjectives, while others are characteristic of nouns. As the table shows, smaller numbers exhibit more adjectival properties, such as agreement with the quantified noun, which means that their syntax is closer to that of a modifier. With these modifying numerals, the plural head of the quantified NP is the noun, and so the notionally plural noun induces plural agreement on the predicate. By contrast, the larger numerals have more noun-like properties, such as distinct singular and plural forms, and the ability to govern agreement on their

			odin 1	*dva* 2	*tri* 3	*pjat'* 5	*sto* 100	*tysjača* 1,000	*million* 1,000,000
Adjective-like	1.	agrees with noun in syntactic number	+	–	–	–	–	–	–
	2.	agrees in case throughout	+	–	–	–	–	–	–
	3.	agrees in gender	+	(+)	–	–	–	–	–
	4.	agrees in animacy	+	+/(–)	+/(–)	–	–	–	–
Noun-like	5.	has semantically independent plural	–	–	–	–	(+)	+	+
	6.	Takes an agreeing determiner	–	–	–	–	–	+	+
	7.	Takes a noun in genitive throughout	–	–	–	–	–	+/-	+

Table 1.1: Repetition of Table 11.3 from Corbett (1983, pg. 228), showing the relatioinship between syntactic behavior and magnitude of the quantifier.

own determiners. Their syntax is therefore closer to a head noun, rather than a noun modifier. When these quantifiers act as heads, rather than the nouns they are quantifying, then they govern predicate agreement, and in that case the agreement they govern is singular.

A similar type of approach can be found in Nichols et al. (1980), who appeal to information structure in order to account for the effect of the relative position of the subject and verb: When subjects precede the verb, plural agreement is more common than when subjects follow the verb. In Nichols et al. (1980)'s account, this variation is related to agreement scope. Minimal agreement scope results in singular verbal morphology, because the verb is simply agreeing with the singular quantifier. Maximal agreement scope results in plural morphology, because the verb is agreeing with the entire notionally-plural subject noun phrase. Information structure is key, because agreement scope is maximal for topics, while for non-topics it is minimal. This means that the relationship between word order and agreement with quantified subjects can be understood as reflecting different agreement scopes for topics and non-topics. Preverbal subjects are topical, which means they condition maximal agreement scope, yielding plural agreement, while postverbal subjects are non-topical, conditioning minimal agreement scope and singular agreement. The heavy influence of word order is not an independent predictor of agreement patterns, but rather it is simply a symptom of the actual factor — namely, the information structure.

The ability of information structure to influence agreement number is also discussed in Lambrecht and Polinsky (1997), who provide the following minimal pair.

(3) Pjat' fil'm-ov pojavi-l-**i**-s' na ekran-ax
Five film-PL.GEN appear-PST-**PL**-REFL on screen-PL.LOC

'(The) five movies were RELEASED.'

(4) Pjat' fil'm-ov pojavi-l-**o**-s' na ekran-ax
Five film-PL.GEN appear-PST-**NEUT.SG**-REFL on screen-PL.LOC
'Five MOVIES were released'/'There were five MOVIES released.'

Example (3) above, showing plural agreement with the quantified subject, is interpreted with what the authors call "predicate focus," corresponding to the neutral topic-comment structure associated with subject-verb word order in Russian. By contrast, example (4) shows the sentence with identical word order, but singular agreement. In this case, the interpretation is one of "sentence focus" — namely, "a pragmatically structured proposition in which both the subject and the predicate are in focus" (pg. 190). This analysis is consistent with that of Nichols et al.: If both the subject and the predicate are in focus in (4), then the subject cannot be a topic, leading to narrow agreement scope and singular agreement on the verb.

These approaches can account for the effect of individual properties on agreement variation, but they are not generalizable to explain other effects, such as subject animacy or verb semantics. Accounts that provide a more unified explanation make use of the idea of notional number, and in particular to the individuation of the quantified subject (Timberlake, 2004). On this view, the effect of the quantifier is not the result of structural properties of the quantifier, or the information structure of the utterance, but rather from the semantics of the quantity being discussed. Larger groups are more difficult to individuate, and for this reason they are more likely to be conceptualized as one unit, reducing the availability of plural agreement. By contrast, smaller numbers allow an easier individuation of the multiple entities, which yields a preference for plural agreement. Non-topics appear post-verbally and with singular agreement not because that particular information status encourages agreement with the quantifier instead of the plural noun phrase, but because non-topics are usually new information, which is harder to individuate than old, familiar entities in the discourse. This approach can be expanded to account for effects of subject animacy, predicate semantics and perhaps even verb aspect, and is attractive because it is so unifying. *Everything* that has been shown to affect verb agreement also has implications for the ability of the speaker to individuate the subject. Further, the effect on individuation aligns with the effect on agreement in each case: More individuated subjects prefer plural agreement, regardless of how that individuation is achieved, be it due to topic status, a larger quantity, animacy, etc.

What the individuation account does *not* provide is a way of calculating the probability of observing plural or singular agreement. Accordingly, one chapter of this dissertation will develop such a tool, building on this unified account of agreement variation from notional individuation by taking into account the simultaneous effects of five factors: animacy, word order, quantifier, verb semantics, and verb tense.

1.3 Theories of agreement production at the word level

In both English and Russian, agreeing forms of verbs are morphologically complex, consisting minimally of a verb root and a suffix. The storage and retrieval of morphologically complex words has been the topic of discussion for several decades, and this section serves to lay out two dominant theories. It is important to note that these theories have tended to focus on words that are complex either due to a derivational process or due to inflection distinguishing present tense from past tense. The review in this section therefore draws on research that is not directly informative as to how agreeing verb forms are processed, and there is at least some evidence that derivationally related forms are processed differently from inflectionally related forms (Laudanna et al., 1992). Ideally, however, these accounts should be generalizable, and the findings presented here will not give any evidence that third-person singular verbs are retrieved or composed differently from other inflected verb forms.

1.3.1 Dual route

One dominant theory of lexical retrieval proposes that there are two routes to the retrieval of inflected forms: retrieval of a complete word form, and assembly from its component morphemes (Baayen and Schreuder, 1999; Pinker, 1999; Pinker and Ullman, 2002a,b). When an inflected form is irregular, it is necessarily stored as a separate unit, and retrieved whole from memory. By contrast, when an inflected form is regular, it can be retrieved either by constructing it according to rule, or by retrieving it from a whole-word representation. Baayen et al. (1997) propose a variant of this theory whereby the dual mechanisms operate in parallel, "racing" against each other, so that the route which selects or constructs the required word form first will be the route by which retrieval is accomplished.

This dual-route model, or variants of it, has been supported by neurolinguistic research. Ullman et al. (2005), for example, observed that left frontal damage impeded patients' ability to produce, read, and judge regular past tense verbs, while temporal or temporal/parietal damage had the reverse effect. Ullman et al. (1997) found similar effects, and further observed that different types of neurodegenerative diseases can selectively impair processing of regular or irregular morphology. Patients with severe Alzheimer's Disease, which damages the ability to learn and recall facts or words, had more difficulty producing irregular verbs, while patients with severe Parkinson's Disease, which can affect grammatical processing, had more difficulty with regular verbs. These dissociations between regular and irregular morphology suggest distinct neural mechanisms for accessing or processing regular and irregular morphology. In particular, difficulties in recalling words are associated with difficulties in processing irregular verbs, while difficulties in grammatical processing lead to difficulties in processing regular morphology. These patterns are expected if irregular forms are stored, while regular forms are constructed by rule (see Ullman, 2001, for a review).

Psycholinguistic evaluations of these models operate on a slightly different logic, building on the fact that high-frequency words are recognized or produced more quickly than low-frequency words. The experimental logic runs as follows: if complex words (e.g., *walked, government*) are stored as decomposed sets of morphemes and assembled as needed, then the frequency with which the base morpheme (*walk, govern*) is used in all of its words combined (*walk, walks, walking, walked* or *government, governed, governor, governance* etc.) will determine how quickly the target word can be assembled. The surface frequency of the target word itself will be irrelevant, because it is not stored as a whole unit. By contrast, if words are stored as whole units, then their retrieval time should be dependent primarily on the frequency of the surface form. The frequency with which the base morpheme is used is irrelevant, because the base morpheme is not being retrieved. On this logic, the distinction between regular and irregular inflection is a subset of the larger distinction between low-frequency words and high frequency words, because irregular verbs tend to be higher frequency than regular verbs (e.g., Baayen and Moscoso del Prado Martín, 2005). According to the dual route model, low frequency inflected words are predicted to be used too infrequently to have whole-word representations, and are thus constructed according to rule. High frequency words, by contrast, do have independent representations, and so it is faster to retrieve them via that route.

These lexical retrieval experiments have found a great deal of evidence in favor of the dual route model. Alegre and Gordon (1999), for example, found that surface frequency effects in English, indicative of whole-word retrieval, can be observed only when the target word's frequency exceeds 6 words per million. With lower-frequency target words, only effects of base frequency can be observed, suggesting decomposed retrieval. Baayen et al. (1997) explored the reaction time differences in identifying Dutch nouns as a function of whether those nouns were low or high frequency, and whether the singular or plural form was more frequent. They found, unsurprisingly, that the plural is identified more slowly than the singular when it is less frequent than the singular, but, more interestingly, the difference was much larger for plurals with low surface frequency than for plurals with high surface frequency. New et al. (2004) found similar results for English and French. Baayen et al. (1997) explained these patterns with a computational model in which the decomposed retrieval route is extremely costly, and only "wins" the race for retrieval when the target word is extremely uncommon. It is this extra processing time required by morphological parsing that explains the added delay in retrieving low frequency plural nouns. Expansions of the theory show that the probability of decomposing a complex word is also affected by properties such as whether a given affix has multiple uses or is productive (Bertram et al., 2000).

1.3.2 Paradigmatic connections

Despite the heavy attention paid to dual-route models of morphological retrieval, more recent work has suggested that lexical storage and retrieval of complex words is a more elaborate matter than proposed by the dual route models. Naming time and lexical decision time is affected not only by the surface frequency of a word and the cumulative frequency of the base, but also by the frequency distributions across related forms, and even the extent to which the frequency distribution for a particular word's inflectional paradigm differs from the average frequency distribution across all that paradigm as a whole (Baayen et al., 2008; Bien et al., 2011; Milin et al., 2009; Tabak et al., 2010). Other work has found that effects of stem frequency and whole word frequency do not appear independently of each other, as would be expected if each effect was the marker of a particular route of access, but instead interact with each other (Baayen et al., 2007; Luke and Christianson, 2011). These findings suggest that the storage or construction of inflectionally related word forms is far more complex than the dual route model, with its choice between whole-word or decomposed representation, allows. Rather, morphologically complex words are stored in an elaborate network of connections, which include information both about the structure of the inflectional paradigm, and the usage patterns indicating how often each form is used with respect to other forms.

A key property of both dual route and more elaborate models of paradigmatic connections is that they are based upon the integration of structure and usage. The heavy emphasis on usage frequency is evident, but the importance of a structural representation is also a fundamental part in these models. Dual route models are built on the fact that the relationship between *adaptor* and *adapt* is different from the relationship between *scream* and *screech*, even though both pairs of words show semantic and formal similarity (Rastle et al., 2000). The more paradigmatically complex models of Baayen and colleagues further posit an ability to recognize the similarity between the inflectional paradigm of a given noun and the inflectional paradigm of all nouns (Milin et al., 2009). The computational methods of these models therefore provide a useful tool for characterizing the probability of selecting a particular form from an inflectional paradigm. In the next section, I lay out how those methods can be implemented.

1.4 Calculating interactions of structure and usage

In the preceding two sections, I summarized existing theories that address how structure and usage patterns influence the production of agreement, both at the sentence-level computation of agreement relations between subject and verb, and the word-level retrieval of a desired inflectional form of the verb. The aim of this dissertation is to infer how these combinations affect production. In order to make this inference, it is necessary to represent the pattern of these combinations by means of some numerical indication of the probability of using a particular structure. This section lays out how these probabilities can be calculated.

1.4.1 The sentence level

Because two different sources of agreement variation — agreement attraction in English and quantified subjects in Russian — will be used in this dissertation, two models will be necessary to calculate the probability of observing a particular agreement form. The English model is the computational implementation of Eberhard et al. (2005)'s Marking and Morphing model, which I laid out in Section 1.2.1. The second is an implementation of the individuation account of Russian agreement variation. Since I build that model and describe its implementation in detail in Chapter 3, I will pass over it here. In this section I will focus on describing the implementation of Marking and Morphing.

According to Marking and Morphing, the probability of observing plural agreement on a verb can be predicted from the plurality specification on the subject noun phrase. This plurality specification, $S(r)$, can be expressed as the sum of two variables: the notional specification $S(n)$ of the subject noun phrase, which is the contribution of the marking stage; and the sum of the contributions of all morphological plurality values $S(m)$ of the nouns contained in the subject, weighted by w, a measure of their syntactic distance from the dominating syntactic node of the subject noun phrase. This is the contribution of the morphing stage. The formula for $S(r)$ is shown below in

Equation 1.1.

$$S(r) = S(n) + \sum_j (w_j \times S(m_j)) \tag{1.1}$$

The notional number specification $S(n)$ can take one of four values. Unambiguously singular noun phrases (e.g., *the key to the cabinet*) have the value 0, while unambiguously plural noun phrases (e.g., *the keys to the cabinets*) have a value of 1. For noun phrases with ambiguous notional number, which include those with distributive interpretations (*the label on the bottles*) or a collective head (*The gang by the motorcycles*), the value of $S(n)$ is a free parameter in the model. Eberhard et al. (2005) determined through a meta-analysis of previous studies of agreement attraction that the best value of $S(n)$ for such ambiguous noun phrases is 0.48. Finally, if the noun phrase has a collective head noun that is interpreted distributively (*The gang on the motorcycles*), then the value of $S(n)$ for ambiguous subjects is doubled, to yield an $S(n)$ of 0.96.

The weights w, which are used to adjust the contribution of the morphological specifications $S(m)$ of the head and local nouns of the subject, are also free parameters determined by Eberhard et al. (2005). For head nouns, the weight is 18.31, while for local nouns it is 1.39.

Finally, the morphological plurality $S(m)$ is calculated by multiplying the morphological specification of a given noun type by its contrastive frequency. Morphological specifications of interest here can take three values: 0 for singular count nouns, 0.07 for singular collective nouns, and 1 for plural nouns, both count and collective. The contrastive frequency for a given noun is equal to the log-transformed summed frequencies of singular and plural noun forms (fq_{SG} and fq_{PL}, respectively), divided by the log-transformed frequency of the plural noun form. The product of these two values — contrastive frequency and morphological specifications — yields the morphological plurality $S(m)$, as shown below in Equation 1.2.

$$S(m) = \text{Specification} \times \frac{log_{10}(\text{fq}_{\text{SG}} + \text{fq}_{\text{PL}})}{log_{10}(\text{fq}_{\text{PL}})} \tag{1.2}$$

As an illustration of how the overall plurality $S(r)$ of a given noun phrase is calculated, consider the subject *the gang with the dangerous rivals*. This subject has a non-distributed meaning, a collective head noun, and a plural local noun. Table 1.2 gives a summary of the values for specifications of notional plurality ($S(n)$), morphological plurality ($S(m)$), and the free parameters w, with the values relevant for this particular subject in bold. Table 1.3 gives the frequency counts of the singular and plural forms of the head and local nouns, taken from the SUBTLEX-US corpus (Brysbaert et al., 2012). The contrastive frequency C_{freq} for each noun is provided in the rightmost column. When this information is combined according to Equations 1.1 and 1.2, the overall plurality $S(r)$ of the subject noun phrase can be calculated as shown in Equation 1.3.

$$\begin{aligned} S(r) &= 0.48 + \left[18.31 \times \left(0.07 \times \frac{log_{10}(1492 + 224)}{log_{10}(224)}\right)\right] + \left[1.39 \times \left(1 \times \frac{log_{10}(120 + 61)}{log_{10}(61)}\right)\right] \\ &= 4.002 \end{aligned} \tag{1.3}$$

Specification	Sg	Pl	Coll Sg	Coll Pl	Coll Sg (distributed)	Weight	Value
$S(n)$	0	1	**0.48**	1	0.96	Head noun (w_H)	**18.31**
$S(m)$	0	**1**	**0.07**	1	0.07	Local noun (w_L)	**1.39**

Table 1.2: Weights and values for notional ($S(n)$) and morphological ($S(m)$) specifications according to Eberhard et al. (2005)'s Marking and Morphing model.

Noun	Sg. fq	Pl. fq	C_{freq}
Head (*gang*)	1492	224	1.38
Local (*rivals*)	120	61	1.26

Table 1.3: Singular and plural frequencies for the head and local nouns in the subject *The gang with the dangerous rivals.*

This $S(r)$ value corresponds to the probability of producing a plural verb by means of a logistic transform, given in Equation 1.4. In this formula b is a free parameter that is added to $S(r)$ in order to ensure that the default value of the transform is 0, corresponding to singular agreement.

$$P(plural) = \frac{1}{1 + e^{-(S(r)+b)}} \tag{1.4}$$

As with the weights for head nouns and local nouns, Eberhard et al. (2005) determined the best-fitting value of b by fitting the model to data from previous experiments on agreement attraction, yielding a value of -3.42. In this way, the predicted probability of observing plural agreement in a sentence with a head noun phrase of *The gang with the dangerous rivals* can be computed as in Equation 1.5 below.

$$\begin{aligned} P(plural) &= \frac{1}{1 + e^{-(4.002+-3.42)}} \\ &= 0.64 \end{aligned} \tag{1.5}$$

In other words, Marking and Morphing predicts that if a hundred people were asked to complete a sentence beginning with *The gang with the dangerous rivals*, 64 of them would use a plural verb.

1.4.2 The word level

In this dissertation, the probability of selecting a particular word form from its paradigm will be calculated in two different ways. The first measure is inflectional entropy (Milin et al., 2009), which has been shown to be predictive of lexical retrieval time in a number of different experiments (Baayen et al., 2007; Bien et al., 2011; Moscoso del Prado Martín et al., 2004; Tabak et al., 2010).

The second is the log-transformed variant ratio of the frequencies of two word forms that can appear in the given sentence. Here, those forms are the singular and plural forms of the inflected verb.

1.4.2.1 Inflectional entropy

Entropy describes the amount of uncertainty associated with some set of options. When those options are members of an inflectional paradigm, entropy describes the uncertainty involved in selecting a form from that paradigm. The inflectional entropy H of a particular verb lexeme X is the negative sum of the log-transformed relative frequencies of the inflected word forms x, each weighted by its relative frequency within the paradigm, $P(x)$ (Equation 1.6).

$$H(X) = -\sum_{x \in X} P(x) log_2 P(x) \tag{1.6}$$

Entropy is affected both by the size of the inflectional paradigm, and by the relative frequency distribution of the paradigm members. The size of the paradigm determines the maximum possible entropy value for the system, such that larger systems have larger possible maximum values. The usage frequencies of the members determine how close to that ceiling a paradigm's entropy can be. When all forms are used equally often, the entropy of the system is at ceiling. As the frequency of one form predominates over the others, then the paradigm's entropy approaches 0.

As an illustration of the effect of paradigm size, consider the verb lexeme *lack*, which is typical of English verbs in that it has four inflectional forms: *lack, lacks, lacking, lacked*. If all four members of the paradigm were chosen equally often, then each would have a relative frequency of 0.25, and the entropy would have the maximum possible value for that paradigm: $H = -(0.25 log_2 0.25 + 0.25 log_2 0.25 + 0.25 log_2 0.25 + 0.25 log_2 0.25) = -4(0.25 \times log_2(0.25)) = 2$. If, by contrast, there were 64 possible forms, which is roughly the case with Russian, then if all forms appeared equally often they would all have a relative frequency of 0.015625. The maximum possible entropy for this system is much higher: $H = -64(0.015625 \times log_2(0.015625)) = 6$.

To illustrate the effect of usage frequencies on inflectional entropy, let us return to the four-form system, which has the maximum possible entropy of 2. Consider now the case where one form is almost always used, with a relative frequency of 0.997, while the other three forms are very rarely used, with a relative frequency of 0.001 each. In this case, the entropy of the system approaches 0, the minimum possible value for that paradigm: $= -\Big(0.997 log_2 0.997 + 3(0.001 log_2 0.001)\Big) = 0.034$.

In actual fact, the part-of-speech tagged SUBTLEX corpus (Brysbaert et al., 2012) indicates that *lack* occurs as a verb in one of its four possible forms a total of 496 times. The individual forms therefore have the following relative frequencies: *lack* $\frac{202}{496} = 0.407$; *lacks* $\frac{106}{406} = 0.214$; *lacking* $\frac{123}{496} = 0.248$; *lacked* $\frac{65}{496} = 0.131$. From these relative frequencies, the inflectional entropy can be calculated as below in Equation 1.7.

		Form	Fq	Log odds	Infl. Entropy
lack	Sg	*lacks*	106	$log(\frac{106}{202}) = -0.64$	1.89
	Pl	*lack*	202	$log(\frac{202}{106}) = 0.64$	
like	Sg	*likes*	3645	$log(\frac{3645}{69662}) = -2.95$	0.59
	Pl	*like*	69662	$log(\frac{69662}{3645}) = 2.95$	

Table 1.4: Measures of word-level probability of singular and plural forms for two verbs with dissimilar frequency distributions.

$$\begin{aligned} H(lack) = &-\Big(0.407log_2(0.407) + 0.214log_2(0.214) + 0.248log_2(0.248) + 0.131log_2(0.131)\Big) \\ =&1.89 \end{aligned} \tag{1.7}$$

By contrast, the verb lexeme *like* has a much more uneven distribution of relative frequencies over its four forms: *like* 0.898, *likes* 0.047, *liking* 0.003, *liked* 0.052. Because the bare verb stem is used so much more often than any other form, the entropy of this system is much lower: 0.59.

1.4.2.2 Variant ratio

Despite evidence that inflectional entropy is predictive of some measures of lexical retrieval of complex forms (Baayen et al., 2007; Bien et al., 2011; Milin et al., 2009; Moscoso del Prado Martín et al., 2004; Tabak et al., 2010), it is limited in one key way: The measure is the same for all members of a given paradigm. This means that it is not possible using inflectional entropy to capture the distinction between the form from a paradigm with the relative frequency of 0.997 and the form from that same paradigm with the relative frequency of 0.001. In order to account for that difference, a second measure of word-level probability of selecting a particular form will be used. In the experiments presented in the following chapters, all words will be embedded in sentences in which it is possible to use either singular or plural agreement. This means that the probability of selecting a particular word form is most useful if it can distinguish between the two particular forms that are of interest: The singular form and the plural form. The measure of paradigmatic probability that captures this distinction will be the log odds of using one form over the other. This measure is the log-transformed ratio of the frequency of the target form of the verb to the frequency of the alternative form. Table 1.4 gives some sample measures for two English verbs with dissimilar frequency distributions. Note that here *log* denotes the natural logarithm.

The variant ratio is indirectly related to inflectional entropy, because entropy increases as frequencies in the paradigm become more similar. A paradigm with maximum entropy will therefore have a log variant ratio of 0 (because the ratio of the frequencies of any two forms will be 1, and *log*(1) = 0) while a paradigm with lower entropy will have log variant ratios with higher absolute values. What log variant ratios contribute over and above inflectional entropy is the principled relation between the log odds of selecting one form and the log odds of selecting its competitor: The variant ratio of the singular form to the plural form is the arithmetic opposite of the variant ratio of the plural form to the singular form. In this way it is possible to represent the probability of selecting a particular inflectional form that is distinct for every form in the paradigm, something that inflectional entropy cannot capture, and which also takes into account the importance of the key alternative form in the paradigm.

1.5 Investigative tool: Pronunciation variation

In order to study how usage and structure interact during speech production, it is necessary to have an investigative tool that indexes production processes and lexical access. In this study, the investigative tool is pronunciation variation, which was chosen by virtue of its observed sensitivity both to structure and to usage, at both the sentence level and the word level.

At the sentence level, pronunciation can vary according to syntactic boundaries. Syllables tend to be longer, for example, near the ends of prosodic phrases; and since the phrases tend to cooccur with syntactic boundaries, this lengthening is thus a cue to syntactic structure (Crystal and House, 1988; Lehiste, 1972; Nagel et al., 1996; Price et al., 1991; Wightman et al., 1992). At the word level, pronunciation varies as a cue to morphological structure. Morphemic segments, such as the plural *-s* in *laps*, tend to be longer than equivalent non-morphemic segments, such as the final *s* in the homophonic *lapse* (Losiewicz, 1992; Smith et al., 2012; Walsh and Parker, 1983). Similarly, syllable rhymes tend to be longer before morpheme boundaries, as with *uff* before the progressive suffix in *puff-ing*, than equivalent segments encoding monomorphemic stems, as in the same *uff* in unsuffixed *puffin* (Sugahara and Turk, 2009).

Yet more noticeably than structure, usage also can be indexed by pronunciation variation. Almost every speech unit has reduced phonetic information in it when it is more probable in the context of the utterance — from segments (van Son and Pols, 2003; van Son et al., 2004) to morphemes (Pluymaekers et al., 2005) to syllables (Aylett and Turk, 2004, 2006) to words (Bell et al., 2009, 2003; Gahl et al., 2012; Gregory et al., 1999; Jurafsky et al., 2001; Schuppler et al., 2012; Tily and Kuperman, 2012, among others) to multi-word contractions (Bybee and Scheibman, 1999; Scheibman, 2000).

Pronunciation is also sensitive to usage patterns at a more abstract level. Gahl and Garnsey (2004), for example, found that the probability of using a particular verbal complement — either a direct object or a sentence complement — affected the pronunciation of the complement-taking verb. When verbs were produced with the more probable complement, they were more likely to have their final [t] or [d] deleted, and verbs that were biased towards sentence-complements showed a reduced effect of phrase-final lengthening before those preferred sentence complements compared to verbs biased towards direct-object complements. Tily et al. (2009) and Kuperman and Bresnan (2012) performed a similar analysis, this time looking at the pronunciation of words in sentences with dative constructions. When speakers were producing the variant that was more probable, words were shorter at the key point in the sentence that disambiguated between a double object construction or a noun phrase followed by a prepositional phrase (Kuperman and Bresnan, 2012).

The studies summarized in this section show that pronunciation is sensitive both to structural and usage properties at the sentence and word level, and also indexes their combination at the sentence level. Evidently, it is an appropriate investigative tool to probe further the combination of structure and usage at both sentence and word level. Accordingly, the focus of this dissertation will be the pronunciation of agreement suffixes, and how that pronunciation varies as a function of the probability of producing the agreement suffixes both at the sentence level and the word level.

1.6 Dissertation overview

1.6.1 English

Chapter 2: Pronunciation with agreement attraction in English The experiment examines the pronunciation of the English present tense singular agreement suffix *-s*, as in *he look-s*. In this experiment, speakers produced sentences with agreement attraction constructions. At the sentence level, probability of observing the singular agreement suffix *-s* in the context of the sentence was determined according to the Marking and Morphing model. As the $S(r)$ value of the subject noun phrase increases, the probability of singular agreement decreases, which means the probability of observing the singular suffix *-s* decreases. At the word level, the probability of observing the singular agreement suffix *-s* was calculated in two ways: as the ratio of the singular frequency to the plural frequency, and as the entropy of the inflectional paradigm. Duration and spectral center of gravity of the *-s* suffix and verb stem were analyzed as a function of both types of probability. The key findings indicate that the duration of the suffix and stem vary according to sentence level probability and word level probability, but in different ways: higher sentence level probability yielded shortening, which I will call “contextual reduction,’ while higher word level probability yielded lengthening, which I will call “paradigmatic enhancement.” There was no effect on spectral center of gravity.

1.6.2 Russian

The next two experiments concern the pronunciation of agreement suffixes in Russian. In this case, rather than using agreement attraction to control the probability of a given agreement suffix, quantified subject noun phrases are used. These allow a much more flexible control over the contextual probability. However, until now there has been no quantitative model according to which it is possible to construct appropriate stimuli. Accordingly, Chapter 3 presents an experiment that allows the building of such a model, while Chapter 4 presents the production experiment itself.

Chapter 3: Russian agreement variation In an Internet-based experiment, Russian participants were shown a sentence frame that systematically varied four factors shown to be influential in affecting number agreement with quantified subject noun phrase. These are order, animacy, verb agentivity, and quantifier. In each frame, the verb was given only in the infinitive, and the participants provided the appropriate inflectional form of the verb. The number and tense of the verb were coded from their results. Verb number was then analyzed via logistic regression modeling, using the four factors in the sentence frame, together with verb tense, as predictors. This quantitative model was then used to construct stimuli for the next experiment, so that those stimuli could be associated with a predicted probability of containing plural agreement.

Chapter 4: Pronunciation with quantified noun phrases in Russian This experiment examines the pronunciation of Russian past tense singular and plural agreement suffixes *-o* and *-i*. Speakers produced sentences with quantified subject noun phrases, which were constructed according to the model built in Chapter 3. The contextual probability of observing a plural or singular agreement suffix at the sentence level was manipulated by means of the quantifier, verb agentivity, word order, and subject animacy of the sentence. The probability of observing a plural or singular agreement suffix at the word level was calculated as in the English experiment in Chapter 2. Duration, F1, and F2 of both suffixes are analyzed as a function of both types of probability. Although there was no effect of either type of probability on suffix duration, vowel quality showed variation on both domains. As the sentence level probability of producing the singular *-o* increased, the difference between *-i* and *-o* on the F1 dimension was reduced (contextual reduction). As the word-level probability of producing a given suffix increased, the difference between *-i* and *-o* on the F2 dimension increased (paradigmatic enhancement).

On the basis of the English and Russian results, I present a theory of how context and paradigm structure interact during production. This theory, which I call Contrast Dependent Pronunciation Variation (CDPV), runs as follows. Pronunciation variation of a morpheme shows up on the phonetic feature that is most important for distinguishing certain salient morphological contrasts. The saliency of these contrasts is driven by the context of the sentence. In the experiments presented here, sentence context restricts the relevant morphological forms to singular and plural agreement. Accordingly, pronunciation varies along the features that distinguish singular and plural forms from each other. In English, that feature is duration. In Russian, that feature is slightly F1, but especially F2.

1.6.3 CDPV and the intersection of structure and usage

Chapter 5: General discussion and conclusions In this final chapter, I summarize my findings and describe how the CDPV provides an understanding of the interaction of structure and usage in speech production. Specifically, I argue that the contextual reduction effects in English are driven by the fact that retrieval of contextually improbable elements is more difficult than retrieval of probable elements, even when probability — a type of usage pattern — is determined on the basis of which structures are likely.

The paradigmatic enhancement is the result of segmental competition between paradigmatically related forms, of the sort proposed by Baese-Berk and Goldrick (2009). A boost in activation is triggered in order to overcome competition from phonologically related forms. This higher eventual activation results in a more extreme articulation. In the present study, agreeing verbs are produced in sentences that permit variation between two paradigmatically related forms. This means that the strongest competitors will be not simply be related phonologically to the target word form, but also related paradigmatically. More frequent forms within that paradigm are stored with a higher resting activation level, which means that their final activation, after receiving the competition-triggered boost, will also be that much higher. This results in a more extreme articulation — the source of paradigmatic enhancement.

The specificity of the enhancement effect, which targets only the features that are salient for distinguishing contextually relevant competitors, is due to the fact that the paradigmatic links between the two competing forms are more complete than the links between forms related solely on the basis of phonological similarity. This is a consequence of the fact that morphological structure is often closely tied to phonological representations, yielding the vast array of morphophonological processes that can be observed in the world's languages. Due to these closer ties between paradigm-mates, the more extreme articulation that is necessary to overcome competition can focus on the features that will most precisely distinguish the target from the competitor.

By my account, structure combines with usage to affect the storage and retrieval of agreeing verb forms, at both the sentence level and the word level. At the sentence level, usage patterns can make it easier to retrieve particular forms, even if those patterns describe the probability of using not the form itself, but a more abstract structure that the form instantiates. This is why singular verbs show phonetic reduction in sentences with a high probability of singular agreement. Even if the verb itself is not probable, it encodes a structure that is probable, and its retrieval and articulation are consequently faster. At the word level, usage patterns affect the storage of individual word forms. The nature of the paradigmatic relations between those word-forms — and crucially the morphophonological relationship — determines how the usage patterns affect the eventual retrieval and articulation of the word. In this way, structure and usage combine at both the sentence and word level in elaborate ways that all converge during the production of subject-verb agreement.

Chapter 2

Pronunciation variation and contextual and paradigmatic probability: Evidence from English

2.1 Introduction

Nothing is ever pronounced the same way twice. The same utterance, spoken by the same person twice, can have vastly different phonetic properties, varying as the speaking register changes (Labov, 1972), or according to conversational partner (Delvaux and Soquet, 2007). One intriguing source of phonetic variation is probability: Depending on how likely a linguistic unit is to be used, the pronunciation of that unit can fluctuate in minute but systematic ways. Yet probability can be measured in a number of different ways, and it can describe multiple different properties. In this chapter I explore the extent to which the nature of probabilistic phonetic variation depends on the type of probability.

Two broad categories of probability — *contextual* and *paradigmatic* — can be illustrated in the table below. In each case, the word of interest is in final position. Contextual probability, describing how likely a given linguistic unit is to be used in some utterance context, is illustrated by the contrast between the columns. In the left-hand column, the final words have high contextual probability: Given the way the utterance begins, the ending word is exactly as expected. In the right-hand column, by contrast, the final word is extremely unexpected in the context of the rest of the utterance.

Paradigmatic	Contextual: High	Contextual: Low
High	The early bird gets the *worm*.	There's more than one way to skin a *worm*.
Low	I don't want to open a can of *worms*.	That's the greatest thing since sliced *worms*!

Table 2.1: Different types of probability.

Paradigmatic probability, which describes how likely a unit is to be selected from some set of related units, is illustrated by the contrast between the rows. Most commonly, paradigms describe morphologically related units, such as inflectional paradigms, for example, or morphological families. English nouns have extremely small inflectional paradigms: Leaving aside the possessive form, the nominal paradigm contains only two members — singular and plural. In the case of *worm*, usage frequencies from the SUBTLEX corpus (Brysbaert et al., 2012) indicate that singular *worm* is more frequent, with 417 uses, than the plural *worms*, with only 320 uses. Thus, the sentences in the top row use singular *worm*, which has high paradigmatic probability, while the sentences in the bottom row use plural *worm*, with low paradigmatic probability.

Previous work on contextual effects on pronunciation has been wide-ranging, exploring units ranging from segments (e.g., van Son and Pols, 2003) to syllables (Aylett and Turk, 2004, 2006) to whole words (e.g. Bell et al., 2009; Lieberman, 1963; Tily and Kuperman, 2012). This research has yielded a robust body of evidence associating higher contextual probability with phonetic reduction of some kind. By contrast, research into paradigmatic effects has focused almost solely on the pronunciation of morphemes, rather than words. Morphemes as a unit have not been addressed systematically in the work on contextual effects, and the effects of paradigmatic probability on pronunciation variation of morphemes have not been as straightforward. In some cases, higher paradigmatic probability is associated with phonetic enhancement (e.g. Kuperman et al., 2007), while in others it appears to condition reduction (Hanique and Ernestus, 2011).

This chapter examines the effect of contextual and paradigmatic probability of subject-verb agreement suffixes, because these suffixes can vary in both domains simultaneously. Since they encode relationships with the subject, which is external to the verb, they can have varying contextual probabilities *independently* of the verb that is used. On the other side, they bear full responsibility for distinguishing one form of the verb from other forms, which means that they also have varying paradigmatic probabilities. The goals of this chapter are, first, to determine whether contextual probability operates upon morphemes in a manner that is consistent with other speech units, and second, to probe the effects of paradigmatic probability on the pronunciation of morphemes. The speech unit selected as a test case will be the singular subject-verb agreement suffix *-s* in English. The key phonetic properties under investigation are the duration of the suffix and its spectral center of gravity, both of which are consistently identified as indicators of phonetic reduction. (e.g., Aylett and Turk, 2004, 2006; Bell et al., 2009; Hanique and Ernestus, 2011; Hanique et al., 2010; Kuperman et al., 2007; van Son and Pols, 2003; van Son et al., 2004).

2.1.1 Contextual probability

Contextual probability describes how likely a given unit is to appear in a particular context. The specific measurement that is used to describe this probability varies from study to study, but the most frequently used measurements make reference to the sequential order of speech units. The earliest research used cloze probabilities, or the rates at which experimental participants produced a particular word in a given linguistic frame (Lieberman, 1963). More modern computational advancements and increasingly large corpora have made it possible to calculate probabilities in more nuanced ways. Joint *n*-gram probability, for example, describes how often *n* linguistic units

— usually words — appear together, normalized by the size of the corpus that provides the frequency data (Bell et al., 2003; Gregory et al., 1999). Mutual information describes how often pairs (or larger sets) of words appear together, compared to how often they might be expect to appear together given their individual frequencies of occurrence (Gregory et al., 1999; Hanique and Ernestus, 2011; Pluymaekers et al., 2005; Tremblay and Tucker, 2011). Conditional n-gram probability describes how often a linguistic unit, such as a word appears in a given context of n other units, divided by how often that context appears with or without the target unit (Bell et al., 2009, 2003; Jurafsky et al., 2001, see also Tremblay and Tucker, 2011, for a variant).

Another way to characterize contextual probability moves away from sequential ordering, and instead focuses on more abstract syntactic structures. In these cases, the measure of contextual probability tends to be more specific to the particular structure at hand. For example, the verbs *confirm*, *announce*, and *believe* all allow either a noun phrase direct object (e.g., *confirm/announce/believe the results*) or a sentence complement (e.g., *confirm/announce/believe the results are correct*). Yet they differ in how likely a given complement is. After *confirm*, the probability of a noun phrase complement is higher than the probability of a sentence complement; after *announce*, the two complements are equally probable, and after *believe* the sentence complement is more probable (Garnsey et al., 1997). The preference for a given complement type is called the *verb bias* (Gahl and Garnsey, 2004; Garnsey et al., 1997), and refers to specifically to this type of variation in argument structure, yet it, too, is a type of contextual probability. Subordinate clauses have a high contextual probability following sentence-complement biased verbs, and noun phrases have a high contextual probability following direct-object biased verbs.

2.1.1.1 Effects of contextual probability on pronunciation

In almost all cases, higher contextual probability has been consistently associated with phonetic reduction on some domain. English words with higher cloze probabilities, for example, tend to have shorter duration (Lieberman, 1963), and words with a higher conditional probabilities have more reduced vowels and shorter duration (Bell et al., 2009, 2003; Jurafsky et al., 2001). Words with higher mutual information values are more likely to have their coronal stops tapped word-medially, or deleted word-finally (Gregory et al., 1999). In Dutch, words with higher bigram conditional probabilities are less likely to epenthesize schwa (Tily and Kuperman, 2012). This reduction effect is not restricted to words. English syllables that have higher unigram, bigram, or trigram conditional probabilities are shorter, with more centralized vowels (Aylett and Turk, 2004, 2006), and some Dutch stems are shorter and have more segments deleted in words with higher mutual information values (Pluymaekers et al., 2005). Even at the level of segments, van Son and Pols (2003) found that segments with higher segmental information content in a word — a measure derived from conditional probability — had shorter duration and lower spectral centers of gravity.

When contextual probability is determined by more abstract structures, rather than linear ordering of linguistic units, the relationship between increasing contextual probability and phonetic reduction holds. Gahl and Garnsey (2004), for example, found that final coronal stops were more likely to be deleted from English verbs when the verbal complement matched the more frequent argument structure of the verb. Demberg et al. (2012) found that words which instantiated more probable syntactic structures had shorter duration, over and above any effect of *n*-gram frequency. Kuperman and Bresnan (2012) found that ditransitive verbs and the immediately following word were shorter when they appeared in sentences that contained the more likely variant of the dative alternation.

This last example is of particular interest, because it illustrates how contextual probability can change as a result of grammatical variation. Grammatical variation is the phenomenon whereby a speaker can use one of multiple structures to express the same thought — as in the case of using a double-object construction with a ditransitive verb, rather than a direct object and a prepositional phrase. Other examples of grammatical variation include alternations between active and passive voice, or the use of singular or plural verb agreement with a collective subject noun. The key property is that grammatical variants are semantically equivalent: the probability of using a particular grammatical structure is insulated to some extent from the conceptual message conveyed by the utterance. This property will therefore be exploited in order to accomplish the first goal of this chapter — namely, to determine whether higher contextual probability conditions phonetic reduction in morphemes. Here, the source of grammatical variation is agreement attraction, or the choice of singular or plural verb agreement with complex subject noun phrases.

2.1.1.2 A special case of grammatical variation: Agreement attraction

In most cases, English verbs agree unambiguously with their subjects. If the subject is third person singular, a present tense verb will have the singular *-s* suffix, as in *he look-**s***. If the subject has any other person or number, the verb will be only the bare stem, as in *they/I/you/we look*. In one particular construction, however, English shows a certain degree of variation in the realization of number agreement. This construction, known as *agreement attraction*, refers to a particular type of speech error first described in depth by Bock and Miller (1991). These errors typically occur in sentences in which the grammatical subject is composed of a head noun that is modified by some structure, such as a prepositional phrase adjunct or a relative clause. The modifying structure contains another noun, called the *local* noun, that differs from the head noun in some feature. Usually the feature is number, with the head noun singular and the local noun plural (e.g., Bock and Cutting, 1992; Eberhard et al., 2005; Gillespie and Pearlmutter, 2013; Haskell et al., 2010; Solomon and Pearlmutter, 2004, among others), but in some languages the head and local nouns can differ in gender, as French, Spanish, and Italian (Antón-Méndez et al., 2002; Franck et al., 2008; Vigliocco et al., 1996b), Dutch (Meyer and Bock, 1999), Russian (Lorimor et al., 2008), and Slovak (Badecker and Kuminiak, 2007). Whatever the feature, agreement attraction occurs when speakers produce a predicate that agrees with the local noun, rather than the head. Some examples of number agreement attraction in English that have been observed in spontaneous speech are given below, with the head noun marked with a subscript *H* and the local noun marked with a subscript *L*.

(5) The description$_H$ of the pictures$_L$ **are** beautiful.

(6) The way$_H$ it's going to be spelled out in the clauses$_L$ **are** going to be different.

(7) The origin$_H$ of those geminate implosives$_L$ that I'm constructing **are** not the result of some morphological process.

The extensive research that has been conducted on agreement attraction has yielded a fairly strong understanding of exactly which properties can affect the rates of agreement attraction. Typically, experiment participants are presented with a sentence preamble, such as *The description of the pictures* ... and then asked to complete the sentence however they like (e.g., Badecker and Kuminiak, 2007; Bock and Miller, 1991; Solomon and Pearlmutter, 2004; Vigliocco et al., 1996b). Other experimental designs involve providing a subject preamble and an adjective or a past participle, and asking participants to create a sentence by combining the two stimuli (e.g., Thornton and MacDonald, 2003; Vigliocco and Nicol, 1998). This approach has the advantage of forcing speakers to use a form of *be*, which agrees in both present and past tense, and thus avoids the problem with free-completion sentences, in which a large number of observations must be discarded because they contain non-agreeing past-tense verbs. In either case, the rates at which speakers produce plural and singular inflected verbs are then analyzed, to determine whether and how much plural agreement rates increase as various factors are manipulated.

One of the earliest factors that was identified through these experiments was syntactic structure. Local nouns have less of an effect when they are more more deeply embedded in the postmodifier (Franck et al., 2002), or when the postmodifier is a clause, rather than a prepositional phrase (Bock and Cutting, 1992, but see Gillespie and Pearlmutter, 2013). Likewise, longer modifying expressions led to higher agreement attraction when the expression was a prepositional phrase, but not when it was a clause (Bock and Cutting, 1992; Bock and Miller, 1991).

Later work established that conceptual numerosity of the subject also plays a role. For example, Vigliocco and colleagues found in Romance languages and Dutch that higher rates of plural agreement occur with subjects like *the label on the bottles*, which can be understood as referring to multiple labels, compared to subjects like *the baby on the blankets*, which can only denote a single baby (Vigliocco et al., 1996a, 1995, 1996b). The effect of this particular sort of conceptual numerosity has not been found in English (Bock and Miller, 1991; Vigliocco et al., 1996a), but Humphreys and Bock (2005) did find a type of conceptual effect that derives from distributive interpretations. Compared to subjects like *the gang near the motorcycles*, which describes the location of a group, people use plural agreement more often with subjects like *the gang on the motorcycles*, whose distributive interpretation emphasizes the fact that the group is composed of separate individuals, each one sitting on a separate motorcycle.

Conceptual effects beyond questions of numerosity can also affect agreement. Thornton and MacDonald (2003) found that local nouns which could be plausible subjects for the verb elicit higher rates of plural agreement than less plausible local nouns. Similarly, Solomon and Pearlmutter (2004) found that the semantic integration of the head and local noun affect rates of attraction, such that more subjects with tighter links between head and local noun elicit higher rates of agreement.

This extensive research has generated not only a wide variety of experimental findings, but also a wide variety of experimental stimuli. These stimuli consist of complex subjects whose propensity to elicit plural agreement is already known. On the basis of these findings, Eberhard et al. (2005) performed a meta-analysis of dozens of previous experiments, in which they developed a computational implementation of their Marking and Morphing model, described in Chapter 1. The current experiment takes advantage of this work by using the same subject preambles that have been developed in previous experiments. For each subject, the probability of observing the singular agreement suffix *-s* on the verb is taken to be the plurality specification calculated by the model, and the pronunciation of the suffix will be analyzed as a function of that plurality specification.

2.1.2 Paradigmatic probability

Paradigmatic probability describes not how likely a linguistic unit is in context, but rather how likely it is to be chosen from some set of forms that are related in some way — through a shared morphological root, for example, or because they are variant pronunciations of the same lexeme. There are a number of different ways of calculating the paradigmatic probability of a particular form. Perhaps the simplest is relative frequency within the paradigm, or the proportion of uses of a particular form in a paradigm out of all uses of any form in the paradigm, a measure that Hanique and Ernestus (2011) used to explore pronunciation variation. Thus, for example, the relative frequency of the word *lacks* in its paradigm is the frequency of all uses of the verb *lacks*, which in the SUBTLEX corpus tagged for part of speech is 106 (Brysbaert et al., 2012), divided by the summed frequency of all words in that paradigm — lacks, 106; lacked, 65; lacking 123; lack, 202. This yields a relative frequency of 106/496, or 0.21. A more commonly used type of relative frequency is a ratio of the frequency of uses of two forms, a measure that is widely used to explain multiple different phenomena. Hay (2001, 2003) drew on this relative frequency measure to account for patterns of semantic opacity and phonetic reduction in words with derivational morphology, and Schuppler et al. (2012) and Hanique et al. (2010) did the same to explore pronunciation variation in inflected Dutch words.

Another variant of relative frequency that is becoming increasingly popular is inflectional entropy. Entropy, described in detail in Chapter 1, takes into account the relative frequency of all forms in the paradigm. It is defined as the negative sum of the log-transformed relative frequency of each inflectional form in the paradigm $P(x)$, weighted by its relative frequency (see Equation 1.6). This measure has frequently appeared as a means of describing the effect of a word's inflectional paradigm to its lexical retrieval time (Baayen et al., 2006; Bien et al., 2011; Milin et al., 2009; Moscoso del Prado Martín et al., 2004; Tabak et al., 2005)

2.1.2.1 Two measures of paradigmatic probability

The current study will examine the effects of two types of paradigmatic probability on the pronunciation of the singular *-s* suffix. The first type will be the inflectional entropy of the verb's paradigm. To date, this measure's effect on pronunciation has not been explored, but it has frequently been implicated in the speed of word recognition. It is often the case that words which are recognized more quickly are also pronounced with some degree of phonetic reduction. High-frequency words, for example, whose advantage over low frequency words in both visual and auditory identification has been known since the 1950s (e.g., Forster and Chambers, 1973; Howes, 1957; Howes and Solomon, 1951; McGinnies et al., 1952), are pronounced with shorter duration (Gahl, 2008). It is therefore reasonable to expect that inflectional entropy, which affects lexical decision and naming reaction time, should also affect pronunciation of those words once they are retrieved. Yet the effect of inflectional entropy on word recognition has been contradictory, making it difficult to predict any corresponding effect on pronunciation. Some work has found shorter naming times and faster, more accurate lexical decision responses to words whose paradigms have higher inflectional entropy (Baayen et al., 2006, 2007; Tabak et al., 2005). Baayen et al. (2007) hypothesize that this is a task-specific kind of effect. Higher inflectional entropy is conditioned by similar usage frequencies of all inflected forms, and these similar usage frequencies can be a strong cue that one of the forms is a real word, thus speeding lexical decision times. This explains the facilitative effects in lexical decision tasks, but it doesn't explain the faster naming times, in which participants simply repeat words as they are displayed on a screen (Baayen et al., 2006). By contrast, Bien et al. (2011) observed an inhibitory effect of inflectional entropy in a position-response association task, in which participants learned to associate a particular word with a position on a computer screen, and then were cued by means of an indicator on one of the screen positions the target word.

The second measure of paradigmatic probability is the relative frequency of the singular verb form (e.g., *looks*) to the plural form (*look*). This measure tends to be examined in work on pronunciation as well as lexical retrieval. Earlier studies used a factorial design, in which the key distinction was whether one form was more frequent or less frequent than another form (Baayen et al., 1997; Hay, 2001, 2003; New et al., 2004). More recent work has started treating relative frequency as a gradient property, measured as some ratio — frequently log-transformed — of the target word frequency to the stem word frequency (e.g., Hanique et al., 2010; Schuppler et al., 2012), or the ratio of the target word frequency to the combined frequency of all possible forms in the paradigm (e.g., Hanique and Ernestus, 2011; Kuperman et al., 2007).

2.1.2.2 Effects on pronunciation

Unlike the research into effects of contextual probability, findings regarding effects of paradigmatic probability on pronunciation have not been consistent. Broadly, there have been two basic sets of results: either paradigmatically probable units are reduced, or else such units are enhanced. Both of these sets of results have been based on relative frequency measures

Investigations into the effect of relative frequency on pronunciation first became widespread in the work of Hay and colleagues, who explored the relationship between stem frequency (e.g., *soft*, *swift*) and derived word frequency (e.g., *softly, swiftly*). These pairs — e.g., *swiftly* and *swift* — belong to the same derivational paradigm, because they all share the same root morpheme. This research revealed that the absolute frequency of derived forms may not be as important in determining lexical retrieval as the relative frequency of a derived form to its base (Hay, 2001; Hay and Baayen, 2001; Hay, 2003). Specifically, Hay and colleagues observed phonetic reduction in morphologically complex words that are frequent with respect to their stems — that is, forms with higher paradigmatic probability. For example, the stem-final *t* is deleted more often in the adverb *swiftly* than in *softly*, which Hay (2003) hypothesizes is a result of the fact that *swiftly* is more frequent with respect to its base *swift* than *softly* is with respect to *soft*. As a result, *swiftly* is therefore retrieved as if it were a monomorphemic form, and phonetically reduced as a result.

This association between whole-word retrieval and phonetic reduction fits well into existing research that has associated monomorphemic words with reduced pronunciation. Sugahara and Turk (2009), for example, found that syllable rhymes tend to be shorter before non-morphemic suffixes (e.g., *uff* in *puffin*) than before similar morphemic suffixes (e.g., *puff-ing*), and multiple studies have found that non-morphemic sound-sequences tend to be shorter and deleted more frequently than similar morphemic suffixes (Guy, 1991; Losiewicz, 1992; Smith et al., 2012; Walsh and Parker, 1983). It is therefore consistent with these findings to propose that forms which are more likely to be retrieved as whole words, rather than morphologically decomposed forms, would be pronounced more like whole words — namely, with phonetic reduction. The work of Hanique and colleagues has provided further evidence in support of this hypothesis, by demonstrating phonetic reduction of both the prefixal *ge-* and suffixal *-t* in Dutch past participles that have higher frequencies within their inflectional paradigms (Hanique and Ernestus, 2011; Hanique et al., 2010).

Yet this particular association between phonetic reduction and paradigmatic probability has not been as robustly supported in the literature as the association between reduction and contextual probability. Notably, Kuperman et al. (2007), in examing Dutch compound words, found that interfixes were *longer* when they were more probable within the paradigm of compound words sharing the same first component. To explain this unexpected result, they propose the *Paradigmatic Signal Enhancement Hypothesis*. According to this account, paradigms of related forms represent "pockets of indeterminacy," from which the selection of one form can be supported to a greater or lesser extent by its frequency of usage within the paradigm. When a form enjoys a large degree of support within its paradim, it is acoustically enhanced as a result. This hypothesis is corroborated by findings from Schuppler et al. (2012), who found that final /*-t*/ in Dutch past tense verbs was pronounced more often if the relative frequency of the past tense form to the stem form was high. Similarly, Hay et al. (2012) found that English plurals which are used more frequently with respect to their singulars have longer *-s* suffixes. In both of these studies, within the paradigm of verb inflections or noun inflections, past tense suffixes and plural suffixes with a higher relative frequency enjoyed greater support, and were therefore phonetically enhanced.

These two accounts — phonetic reduction springing from whole-word retrieval, and phonetic enhancement springing from paradigmatic support — are directly contradictory to each other, as indeed they must be to explain seemingly contradictory data. It is in part this contradiction that motivates the second goal of this chapter — namely, to explore the nature of effect of paradigmatic probability on pronunciation. To investigate these effects, the pronunciation of the singular *-s* suffix will be analyzed as a function both of the inflectional entropy of the verb and the relative frequency of the singular form to the plural form.

2.1.3 Key questions

The two goals of this chapter are, first, to determine whether increased contextual probability affects morpheme pronunciation in the same way it affects other speech units, and, second, to determine what effect paradigmatic probability has on pronunciation variation. To address these goals, the current study will be structured around two key questions.

Question 1: *Are contextually probable singular agreement suffixes phonetically reduced?*

By answering this question, it is possible to fill a gap in the research into effects of contextual probability — namely, the effect on morphemes. Pluymaekers et al. (2005) come close, by addressing the behavior of the stem and the suffix *lijk* of seven Dutch adverbs indipendently of each other; but even in that study the pronunciation variation in the morphemes was conditioned by the probability of the entire word, not of the morpheme itself. One possible reason that morphemes have been thus neglected is because it is extremely difficult to calculate the contextual probability of a morpheme without confounding the contextual measurement with paradigmatic measurements. To illustrate: the contextual probability of seeing a suffix like *-al* given a preceding stem *nation*, (as in *national*) is necessarily determined by the number of other words containing the same stem *nation*. These other words, however, form the morphological family of the word *national*, and so the measure of contextual probability of the suffix *-al* can be derived directly from the word's paradigmatic probability. By focusing on subject-verb agreement morphemes in English, therefore, the contextual probability of an agreement suffix can be calculated independently of the verb stem, because the property that determines how likely this suffix is in the context of the sentence is the structure of the subject. When the subject has a low plurality specification, singular verbs are very probable, and because that singular agreement is encoded by the suffix *-s*, the suffix itself is likewise very probable. This is true regardless of the identity of the verb stem, which makes it possible to determine the effect of contextual probability on pronunciation independently of paradigmatic probability.

Question 2: *Do relative frequency and inflectional entropy have distinct effects on lexical retrieval?*

Recall that higher relative frequency has been found to be associated both with phonetic enhancement (Kuperman et al., 2007; Schuppler et al., 2012) and reduction (Hanique and Ernestus, 2011; Hanique et al., 2010; Hay, 2003). Similarly, higher inflectional entropy has been found to condition faster reaction time in lexical decision or naming tasks (Baayen et al., 2006, 2007; Tabak et al., 2005) and also slower reaction time (Baayen et al., 2008; Bien et al., 2011). These contradictions make it difficult to predict the nature of the effect of inflectional entropy on pronunciation.

One possibility is that relative frequency and inflectional entropy have distinct effects on pronunciation. If it is the case that inflected word forms are stored with multiple links to related forms from the same paradigm, then it would be more difficult to navigate a paradigm with higher entropy. The difficulty in selecting a target form from such a paradigm is what yields the slower naming time observed by Bien et al. (2011), and that same delay could also manifest itself in slower articulation and phonetic enhancement. This effect would apply to all forms within the paradigm, regardless of their relative frequency. By contrast, relative frequency applies to a distinct form within the paradigm. As Hay (2003) proposes, forms which are retrieved more easily, perhaps with the help of whole-word representations, are also produced with a reduced pronunciation.

This dissociation can also go the other way. If it is the case that paradigms with high entropy are quickly named, as observed by Baayen et al. (2006), then the faster naming time might result in reduced pronunciation. By contrast, consistent with the Paradigmatic Signal Enhancement Hypothesis (Kuperman et al., 2007), forms with high relative frequency within their paradigms would be phonetically enhanced.

The key explanatory power of proposing a dissociation between effects of entropy and relative frequency is that the two are not entirely independent. Entropy increases as relative frequencies in a paradigm become more similar. For that reason, as the relative frequency of an uncommon form increases, entropy also increases. This means that some findings which seem to show inhibitory effects of entropy might actually be showing inhibitory effects of relative frequency, and vice versa. By clearly distinguishing between the two measures of paradigmatic probability, it might be possible to account for the previous contradictory findings. The following analysis of *-s* pronunciation will therefore orthogonalize these two measures in order to find whether they have distinct effects on pronunciation, and, if they do, to determine the direction of those effects.

2.2 Methods

The experiment presented here was designed to answer the two key questions. In this experiment, speakers produced sentences containing agreement attraction constructions and a variety of singular verbs. The stimuli were designed to allow the *-s* suffix to vary in both types of probability. Contextually, the probability of observing a singular *-s* suffix depended on the probability of observing singular agreement in the sentence. Paradigmatically, the probability of observing the *-s* suffix depended on the particular verb that was used. In this experiment, there were 53 distinct verbs, for which measures of both inflectional entropy and singular/plural relative frequency were calculated.

2.2.1 Materials

2.2.1.1 Verbs

Fifty-three distinct one-syllable verbs were selected, each one ending in either /p/ or /k/ to ensure that the singular agreement allomorph was always *-s*. The frequencies associated with each verb and the other members of the verb's paradigm were taken from the part-of-speech tagged SUBTLEX corpus (Brysbaert et al., 2012), to ensure that the usage frequencies were describing only the verbal usages of each word. The inflectional entropies for the verbs ranged from 0.59 to 2.05, and the singular/plural relative frequencies of each verb were log-transformed, and ranged from -6.174 to 2.305.

For most verbs, the stem, which is also the plural form, has a much higher frequency than the singular form, largely because this form has so many more uses. Beyond the present tense plural, the bare stem is used for imperatives, infinitives, and also for every present tense form except the third person singular. This means that when the singular/stem relative frequency increases, it approaches 1: the frequencies of the singular form and the plural form become more similar,

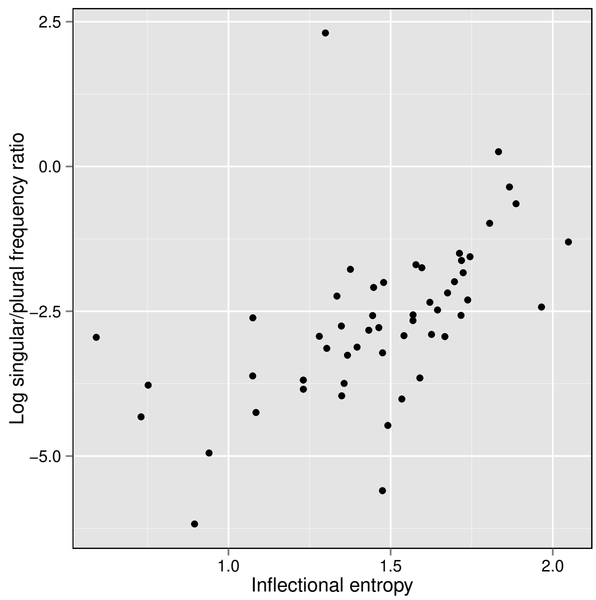


Figure 2.1: Entropy of each verb plotted against the log singular/stem relative frequency.

rather than less similar. As a result, the inflectional entropy of the system as a whole increases, which means that the singular/stem ratios of these verbs are strongly correlated with the verbs' inflectional entropies (Spearman's $\rho = 0.51$), as illustrated in Figure 2.1. These measures were therefore orthogonalized, as described in Section 2.2.5.3.

2.2.1.2 Sentence preambles

Each stimulus list contained 60 critical sentences with grammatical subjects taken from the literature on agreement attraction. In order for the plurality specification of the subjects to vary as much as possible, six categories of subject types from previous literature were selected, such that the average rate of singular agreement that had been observed varied widely. These categories are summarized in Table 2.2. In each category, the grammatical subjects were taken from the previous studies with as little change as possible. Examples of such subjects in each category are provided, along with the rate at which these preambles were reported to elicit plural agreement.

Singular-Singular (Sg-Sg) and Singular-Plural (Sg-Pl) Twenty-four sentences in each list contained subjects with non-collective head and local nouns — meaning that each head and local noun denoted an individual, rather than a group. Half the local nouns were singular, while the other half were plural, counterbalanced across stimulus lists. These subjects were taken from the stimuli used in Gillespie and Pearlmutter (2013), Experiment 1, and Thornton and MacDonald (2003), Experiment 1. The subjects from Gillespie and Pearlmutter (2013) all used the preposition *with* to introduce the modifier containing the local noun, while those from Thornton and MacDonald

Syntactic context	Example	P(sg.)
Sg-Sg	The pizza with the missing slice ...	0.99
Sg-Pl	The pizza with the missing slices ...	0.87
Coll-Sg	The gang with the dangerous rival ...	0.76
Sg-CollPl	The strength of the armies ...	0.67
Coll-Pl	The gang with the dangerous rivals...	0.48
D-CollPl	The jury in the folding chairs...	0.25

Table 2.2: Sample subjects for the critical stimuli and previously reported rates of plural agreement.

(2003) varied the prepositions. Rates of plural agreement for both the Sg-Sg sentences and Sg-Pl are given in parentheses.

(8) The pizza with the missing slice(s) ... (0.004/0.084) (Gillespie and Pearlmutter, 2013, Epx. 1)

(9) The song by the folk singer(s) ... (0.019/0.182) (Thornton and MacDonald, 2003, Exp. 1)

Collective-Singular (Coll-Sg) and Collective-Plural (Coll-Pl) Eighteen sentences in each list contained subjects with collective head nouns, denoting groups, and non-collective local nouns. Half the local nouns were singular, while half were plural, counterbalanced across stimulus lists. These subjects were constructed from the subject preambles used in Bock et al. (2006), Experiment 2. Although it may seem counterintuitive that adding a singular local noun would raise rates of plural agreement over the extremely low baseline reported for simple collective subjects, subject preambles of the sort given below in (10) (e.g., *The gang with the dangerous rival*) did appear with plural agreement rates substantially higher than the rates in the previous two categories. For this reason these subject preambles were used to construct experimental stimuli. Some examples are given below, followed by the plural agreement rates reported for the singular and plural attractors.

(10) a. The gang with the dangerous rival(s) ...
b. The choir for the church service(s) ... (0.238/0.518) (Bock et al., 2006, Exp. 2)

Singular-Collective Plural (Sg-CollPl) A further 9 subjects in each stimulus list consisted of singular head nouns and plural collective local nouns. These were taken from Bock and Eberhard (1993), Experiment 4, and Bock et al. (2006), Experiment 3. These preambles were extracted from two different sources because the stimuli used by Bock and Eberhard (1993) could not all combine plausibly with the selected verbs, and so they were supplemented by some of the preambles used in Bock et al. (2006). These supplementary preambles were not identical to those used by Bock et al. (2006), because the original materials had singular head nouns with singular collective local nouns. Here, the local noun was pluralized, and the expected agreement rate for these subject preambles was assumed to be the same rate reported in Bock and Eberhard (1993). Examples of these sentences are given below in (11–12).

(11) The strength of the armies . . . (0.33) (Bock and Eberhard, 1993, Exp. 4)

(12) The future of the Russian brigades (Bock et al., 2006, Exp. 3)

Distributed Collective-Plural (D-CollPl) The remaining 9 subjects in each stimulus list were taken from Humphreys and Bock (2005). Superficially, these subjects resemble those in the Coll-Pl condition described above, but there is an important difference in interpretation. Whereas the subjects in the Coll-Pl describe a collective group of people modified by some property, the subjects in this category encourage the interpretation that it is not the group for which a property holds true, but rather each individual within the group. Thus, in (13a), for example, it is understood that each member of the jury is in a separate folding chair, and in (13b), the members of the class are distributed among several lunch tables. As Humphreys and Bock (2005) found, this significantly increases rates of plural agreement above the rates found with similar subjects that have non-distributed interpretations.

(13) a. The jury in the folding chairs . . .
b. The class at the lunch tables . . . (0.745) (Humphreys and Bock, 2005)

The six categories of subject type are summarized in Table 2.2, in ascending order of previously observed rates of plural agreement. For the Sg-Sg and Sg-Pl categories, the observed rates in each category from Gillespie and Pearlmutter (2013) (example 8) and Thornton and MacDonald (2003) (example 9), have been averaged (G&P Sg-Sg: 0.004, Sg-Pl: 0.084; T&M Sg-Sg 0.019, Sg-Pl 0.182): . These singular-agreement rates represent the probability of observing a singular agreement *-s* suffix in sentences containing each of the six subject categories.

Stimuli sentences were created by combining these subjects with the verbs described in Section 2.2.1.1, which were followed by an additional 7-12 syllables of linguistic material. Each post-verb sequence started with either a vowel or a stop, in order to make it easier to segment the *-s* agreement suffix on the verb. At the end of each stimulus was a blank space, which participants were asked to complete by selecting one of two provided completions. In this way, speakers were required to process the sentence's structure and meaning, in order to select the best-fitting completion, but their attention was directed away from the subject and verb. An example stimulus is given below in (14).

(14) The pizza with the missing slices looks unappetizing in the ______
morning basket

2.2.2 Design

Four experimental lists were designed, combining the different categories of subject type as follows. For the Sg-Sg and Sg-Pl conditions, 36 sentences with subjects drawn from Gillespie and Pearlmutter (2013) and Thornton and MacDonald (2003) were created with both a Sg-Sg form and a Sg-Pl form. These sentences were then subdivided into groups of 12 and rotated through the four lists and two conditions. In this way, each list had 12 Sg-Sg and 12 Sg-Pl sentences, each sentence appeared in at least one list in both the Sg-Sg and Sg-Pl condition, and no list contained the same sentence in more than one condition.

For the Coll-Sg and Coll-Pl conditions, 18 sentences were created with the subjects drawn from Bock et al. (2006). Each sentence appeared in both the Coll-Sg and Coll-Pl conditions. These sentences were divided into two groups of 9 and rotated through the four stimulus lists and two conditions. In this way, each list had 9 Coll-Sg and 9 Coll-Pl sentences, each sentence appeared in two stimulus lists in both the Coll-Sg and Coll-Pl conditions, and no list contained the same sentence in more than one condition.

For the Sg-CollPl condition, 18 sentences were created with subjects drawn from Bock and Eberhard (1993) and Bock et al. (2006). These sentences were divided into two groups of 9. Half the stimulus lists contained the first group of 9, and the other stimulus lists contained the second group. The same design governed the distribution of the D-CollPl condition. The 18 sentences used in Humphreys and Bock (2005) were divided into two groups of 9 and divided among the stimulus lists.

In this way, a total of 90 sentences were created (provided in Appendix A), but each stimulus list contained only 60 critical stimuli. In each list, these sentences were randomly interspersed with 60 filler sentences, which were of two types. The first type of filler had similar syntactic complexity to the critical sentences, but the gaps were in the middle or the beginning of each sentence. The second type of filler consisted of questions with no gaps to be filled in. Subjects were instructed simply to read these questions as smoothly as possible.

2.2.3 Participants

Forty speakers from the University of California, Berkeley (17 male, 23 female), were recruited to participate in the experiment through the phonology lab subject mailing list and posters. All were self-identified native English speakers.

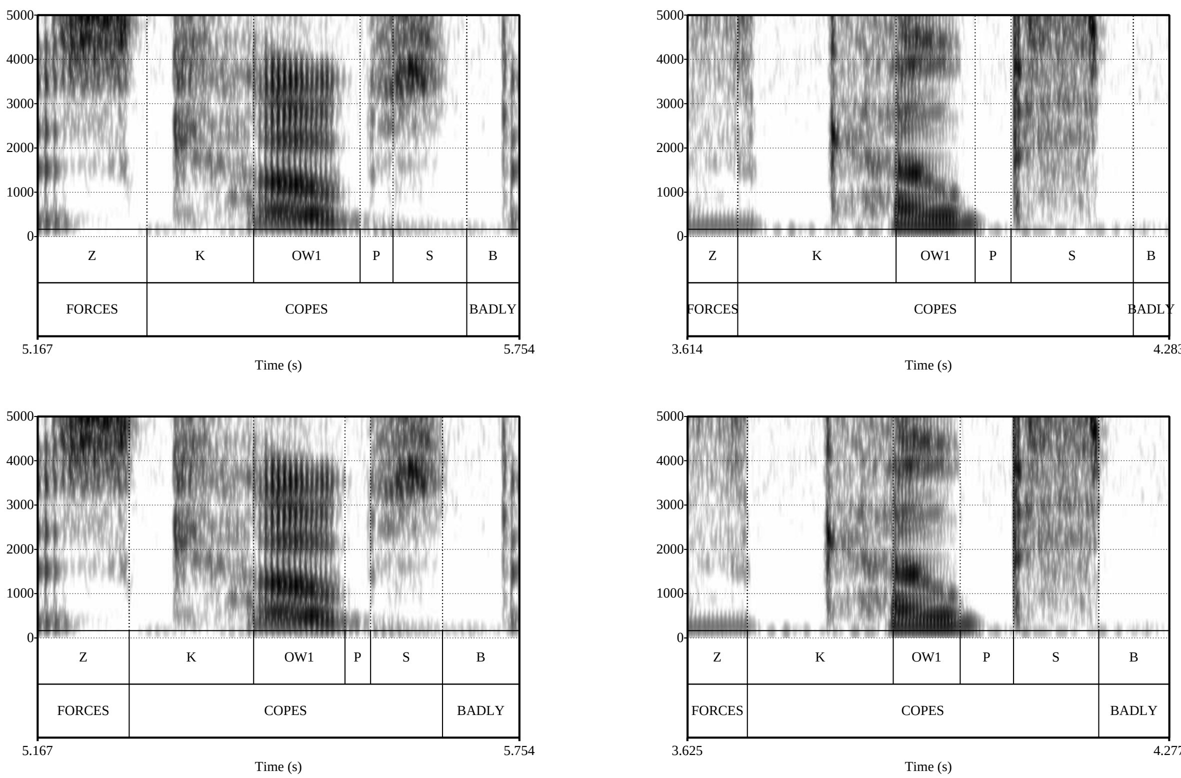
2.2.4 Procedure

Participants were tested individually in a sound-dampened room on a Windows 7 PC running the OpenSesame experimental software (Mathôt et al., 2012). The 120-sentence stimulus list was presented in four blocks of 30 sentences each, and participants were encouraged to take a break between blocks. Sentences were presented in center-aligned black 18-point mono-spaced font on a white background, with line-wrapping after 50 characters. Participants were instructed to read each sentence through and select a completion before speaking, in order to ensure fluency of production. The procedure began with three practice sentences: two fill-in-the-blank, and one question. This allowed participants to get a sense of the allotted time, and ask questions before beginning the experiment. Initially participants were allowed 10 seconds for each sentence, but after the first 13 participants had completed the experiment, it became clear that ten seconds were not sufficient, and so the remaining 27 participants were allotted 15 seconds per sentence. The entire experiment took about 45 minutes to complete, and participants were compensated $7.50.

2.2.5 Analysis

2.2.5.1 Segmentation and coding

Sentences were segmented in two stages. First, they were run through an automatic forced aligner to create unbiased segmentation files matching phonetic segments to acoustic intervals. As it turned out, however, the automatic aligner was not sufficiently accurate, and so the alignment of the verb onset, the singular -*s* suffix onset, and the -*s* offset were hand-corrected. An example of a spectrogram before and after hand-correction is provided in Figure 2.2 to illustrate the need for more precise segmentation than what the automatic aligner could provide.



(a) Alignment produced by the automatic aligner with consistent rightward shift (top), and correction (bottom).

(b) Alignment produced by the automatic aligner with accurate -*s* onset but inaccurate offset (top) and correction (bottom).

Figure 2.2: Examples of alignments produced by the automatic aligner (top row) and after hand-correction (bottom row). Note that the aligner errors are not consistently inaccurate, and so could not be repaired through any systematic adjustment to boundary locations.

Because stop bursts were not always present or easily distinguishable from fricative onsets, all -*s* suffix onsets were marked at the first zero-crossing preceding the noise that marked the closure release of the preceding /k/ or /p/. The offsets of the -*s* suffixes were marked at whichever of the following cues occurred first: (i) the offset of frication noise before a following silence; or (ii) the first zero crossing before the onset of periodic voicing of a following vowel.

From these corrected alignments, the following information was extracted: The duration of the sentence from onset of speech to the verb; the duration of any pause preceding the verb; the duration of the verb stem; the duration of the suffix, and the duration of any pause following the verb.

In addition to segmentation, each sentence was marked for any of the following properties:

- Whether the speaker produced a plural verb instead of a singular one.
- Whether there was a disfluency in the sentence. In most cases this disfluency took the form of a hesitation before the final gap, indicating that speakers had not yet chosen their completion before they began speaking, or else that they had forgotten it before they reached the position of the gap.
- Whether there was an error in the sentence, and if so, whether it was corrected, and whether it occurred before or after the verb.
- Whether the speaker ran into difficulty and restarted the sentence
- Whether the speaker ran out of time to complete the sentence
- Whether the speaker was still producing the preceding sentence when the trial began
- Whether the speaker produced a plural verb instead of a singular one
- Whether the sentence should be discarded entirely. These cases usually involved a pronunciation error immediately at the verb, an inaccurate production of the subject preamble, or else prosodic evidence that the speaker had fallen victim to a garden-path or otherwise not fully understood the sentence.

2.2.5.2 Coding probabilistic predictors

Contextual probability was coded according to Eberhard et al. (2005)'s Marking and Morphing model, described in Chapter 1. For each sentence, the $S(r)$ value was calculated. These values represent the contributions both of notional plurality, which accounts for the distinction between collective head nouns with a distributive or a non-distributive interpretation, and also the morphological properties of the nouns that make up the subject, which are affected by frequency distributions of singular and plural forms. They therefore represent a more fine-grained representation of the plurality value of each subject, beyond what can be inferred from the probability associated with the six bins of subject type shown in Table 2.2. These values, represented by the variable **sap**, were used as the predictor of contextual probability of plural agreement. Therefore, as **sap** decreases, the probability of singular agreement increases.

Paradigmatic probability was represented in two ways: the log-transformed ratio of frequency of the singular to the frequency of the plural form (**lgRatio**), and the inflectional entropy of the verb (**entropy**). As shown in Figure 2.1, these measures are correlated, and so they were orthogonalized in two parallel ways. In the first case, **entropy** was entered as the outcome variable in a simple linear regression model, with **lgRatio** as the only predictor. The residuals of this model (**entropyResid**) represent the variability in inflectional entropy that cannot be accounted for by relative frequency. As a complement to this orthogonalization procedure, a reverse model was built, in which **lgRatio** was predicted from **entropy**. The residuals of this models, **ratioResid**, represent the variability in relative frequency that cannot be accounted for by inflectional entropy.

2.2.5.3 Model-building procedure

Three sets of mixed-effects regression models were built, all according to the same procedure. The first set of models analyzed log-duration of the *-s* suffix, the second analyzed log-duration of the stem, and the third analyzed spectral center of gravity (CoG) of the suffix. In all cases, the models contained random intercepts for participant and verb, and fixed effects were included through forward addition, starting with the control predictors listed in Table 2.3. This method of model entry makes it possible to ensure that each new predictor warrants the increase in model complexity by contributing to model fit. New predictors were retained if they significantly improved model fit, as established by a log-likelihood ratio test and a drop in the Akaike Information Criterion (AIC).

When the predictors in the control model were determined, the three probabilistic variables of interest were added. Each key predictor was considered first as the sole addition to the model, and then as the final addition to a model containing other key predictors. Thus, for example, **lgRatio** was first evaluated by adding it to a model containing only control predictors. If it improved the fit of the model, then another key predictor was added.

The data did not always permit model convergence if more than one random slope was included, so it was impossible to keep the random effects structure maximal (Barr et al., 2013). For this reason, after the fixed effects were determined, each fixed effect was considered as a random slope both for participant and for verb. If more than one random slope improved the model fit according to a log-likelihood ratio test, then as many slopes as possible were added, while still permitting model convergence. All slopes were included in reverse order of the magnitude of the drop in AIC that they yielded, in order to ensure that those with the strongest effect were included first. All slopes were added after key predictors (**lgRatio**, **entropy**, or **sap**) in order to allow for the case that the best slopes might in fact be one of those predictors.

Predictor	Description	Factor levels
Control predictors		
precPause	The presence of a preceding pause	*no, yes*
folPause	The presence of a following pause	*no, yes*
rate	Speaking rate in syllables per second of the sentence preamble, up to the verb (**rate**)[1]	
lgLocalRate	Log-transformed local speaking rate of the verb stem, in segments per second	
lgCumFq	Log-transformed cumulative frequency of all forms of the verb	

[1]Although speaking rate was bounded, such that it could not have a negative value, it was nonetheless roughly normally distributed. Log-transforming it would have removed the boundedness problem, but introduced a skew in the distribution of the transformed rate, and thus the data were not transformed. This trade-off was not an issue for any of the other rates or duration measurements, because log-transforming their distributions did not result in non-normality.

Predictor	Description	Factor levels
meanPhon	Mean positional monophone of the verb (Vitevitch and Luce, 2004)	
meanBiPhon	Mean biphone probability of the verb (Vitevitch and Luce, 2004)	
position	Trial number	
nPhon	Number of phonemes in the verb stem	
folMoa	Manner of articulation of the first sound in the following word	*stop (s), vowel (v)*
folPoa	Place of articulation of the first sound in the following word	*p, t, k, v*
Key predictors		
sap	Plurality specification of sentence subject	
lgRatio	Log-transformed relative frequency of singular form to plural form	
entropy	Inflectional entropy of the verb	
ratioResid	Log-transformed relative frequency residualized on entropy	
entropyResid	Inflectional entropy residualized on relative frequency	

Table 2.3: Predictors and their abbreviations used during the model-building procedure. Control predictors are in the lower block, while key predictors are in the lower block.

2.3 Results

2.3.1 Data pruning

Of the forty participants who completed the experiment, two turned out to be insufficiently fluent in English to complete the sentences in the allotted time, and a third turned out to be a speaker of New Zealand English. Since speakers of other dialects of English do not treat collective subjects in the same way as American English speakers (Bock et al., 2006), the estimates of the contextual probability of plural agreement were not necessarily valid for this speaker. Responses from all three of these participants were therefore excluded from analysis.

Of the remaining 2220 sentences, an additional 128 were discarded for excessive disfluency or prosodic evidence that the participant did not fully understand the stimulus before beginning to speak; 21 more for an inaccurate production of the preamble; and a further 27 because the participant produced a plural verb, rather than a singular verb. Five more outlier observations with excessively extreme stem durations or *-s* durations were discarded. Finally, all observations containing the verb *gasps* were discarded, for two reasons. First, the final *-sps* sequence was frequently reduced by speakers, to the extent that the [p] was represented only by a slight rounding in the middle of a long final [s]. This made it extremely difficult to separate the suffix [s] from the stem [s] accurately. Second, the verb *gasps* had an unusually high singular/stem ratio in the SUBTLEX corpus — 10.02 — which is over seven times greater than the next highest ratio of 1.89, for the verb *trips*. This suggested that some property of its usage in that corpus was not parallel to the usage of the other verbs, due, perhaps, to the heavy use in subtitles of verbs like *gasps* or *sobs* in the singular form to spell out non-linguistic utterances.[2] For these reasons it was determined that *gasps* could not be accurately analyzed, both because of the uncertainty associated with its segmentation, and the uncertainty associated with its usage frequency, and thus all observations were discarded.

The remaining 2001 observations were retained for analysis.

2.3.2 Suffix durations

2.3.2.1 Model summary

Log durations of the *-s* suffix were analyzed with a mixed effects linear regression model built according to the procedure outlined in Section 2.2.5.3. The control predictors that improved model fit were **rate**, **lgLocalRate**, **folPause**, **precPause**, **meanBiPhon**, **meanPhon**, and **folPoa**. Unsurprisingly, as local speaking rate of the verb stem increased, suffix duration decreased. The presence of a following pause yielded substantial lengthening, while a preceding pause yielded more moderate lengthening. Higher biphone probability conditioned dramatic lengthening, while higher positional monophone probability had the reverse effect, to a much lesser degree. Finally, the place of articulation of the first sound of the following word interacted with global speaking rate of the entire sentence up to the verb. With following [t] or [k], an increase in global speaking rate yielded the expected reduction in suffix duration. With a following [p] or vowel, however, there was no such effect. These findings are detailed in the top block of Table 2.4, and illustrated in a partial effects plot in Figure 2.3.

The key predictors, **sap**, **lgRatio**, and **entropy**, were each tested to see how they improved the fit of the control model. There was no improvement to be had by adding **entropy**, whether it was added directly to the control model, or residualized and added to the model already including relative frequency. Increased contextual probability trended towards decreasing suffix duration ($\beta = -0.006, SE(\beta) = -0.003, t = -1.87$), but its contribution to model fit did not quite reach significance ($\chi^2 = 3.5, p = 0.06$). Increased relative frequency, however, consistently yielded longer suffix duration ($\beta = 0.02, SE(\beta) = 0.01, t = 1.9$), and this effect improved model fit both regardless of whether **lgRatio** was added directly to the control model or when it was residualized and added to the model already containing **entropy** (in both cases, $\chi^2 = 3.9, p < 0.05$). The partial effect of relative frequency is included in the middle right panel of Figure 2.3. Because **lgRatio** was the only term that produced a significant improvement over the control model, this fixed effects structure — the control model plus **lgRatio** — was used for all further analyses.

[2]For example, the verb *sobs* has a singular/stem ratio of 5.1 in the SUBTLEX corpus, which, although not as high as *gasps*, is still dramatically greater than *trips*.

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p	R^2 (ΔR^2)
0	**Intercept**	-1.30	0.18	-7.19	-82			0
1	**rate**	-0.09	0.04	-2.31	-114 (-32)	34.02	< 0.001	0.023 (0.023)
2	**lgLocalRate**	-0.23	0.03	-8.62	-216 (-102)	103.62	< 0.001	0.103 (0.080)
3	**folPause**				-378 (-162)	163.77	< 0.001	0.154 (0.051)
	yes	0.28	0.02	12.92				
4	**precPause**				-382 (-4)	6.19	< 0.05	0.160 (0.007)
	yes	0.03	0.01	2.24				
5	**meanBiPhon**	28.74	8.00	3.61	-386 (-4)	6.26	< 0.05	0.171 (0.010)
6	**meanPhon**	-2.32	1.60	-1.45	-391 (-5)	7.00	< 0.01	0.180 (0.009)
7	**folPoa**				-414 (-23)	29.40	< 0.001	0.202 (0.022)
	p	-0.51	0.17	-2.92				
	t	0.14	0.19	0.73				
	v	-0.34	0.16	-2.07				
8	**rate:folPoa**				-429 (-15)	20.50	< 0.001	0.211 (0.009)
	p	0.10	0.04	2.41				
	t	-0.02	0.04	-0.39				
	v	0.06	0.04	1.58				
9a	**lgRatio**	0.02	0.01	1.92	-430 (-1)	3.91	< 0.05	0.213 (0.002)
10a	**entropyResid**	-0.03	0.04	-0.68	-429 (+1)	0.55	0.46	0.213 (0.000)
9b	**entropy**	0.02	0.03	0.71	-427 (+3)	0.54	0.46	0.211 (0.000)
10b	**ratioResid**	0.02	0.01	1.90	-429 (-2)	3.91	< 0.05	0.213 (0.002)
9c	**sap**	-0.006	-0.003	-1.87	-430 (0)	3.5	0.06	0.213 (0.002)
10c	**lgRatio**	0.02	0.01	1.69	-432 (-2)	3.09	0.08	0.215 (0.002)
Refit 8+slope	**folPause** by **verb**				-944			0.259
Refit 9a+slope	**lgRatio**	0.02	0.01	2.03	-947 (-3)	4.12	< 0.05	0.261 (0.002)

Table 2.4: Duration model of the suffix. Blocks 9a-10c indicate the effect of adding each predictor to the control model (line 8). The final block represents the effect of adding **lgRatio** when the control model and Model 9a were fit with random slopes and without outlier observations. ΔR^2 is calculated before rounding.

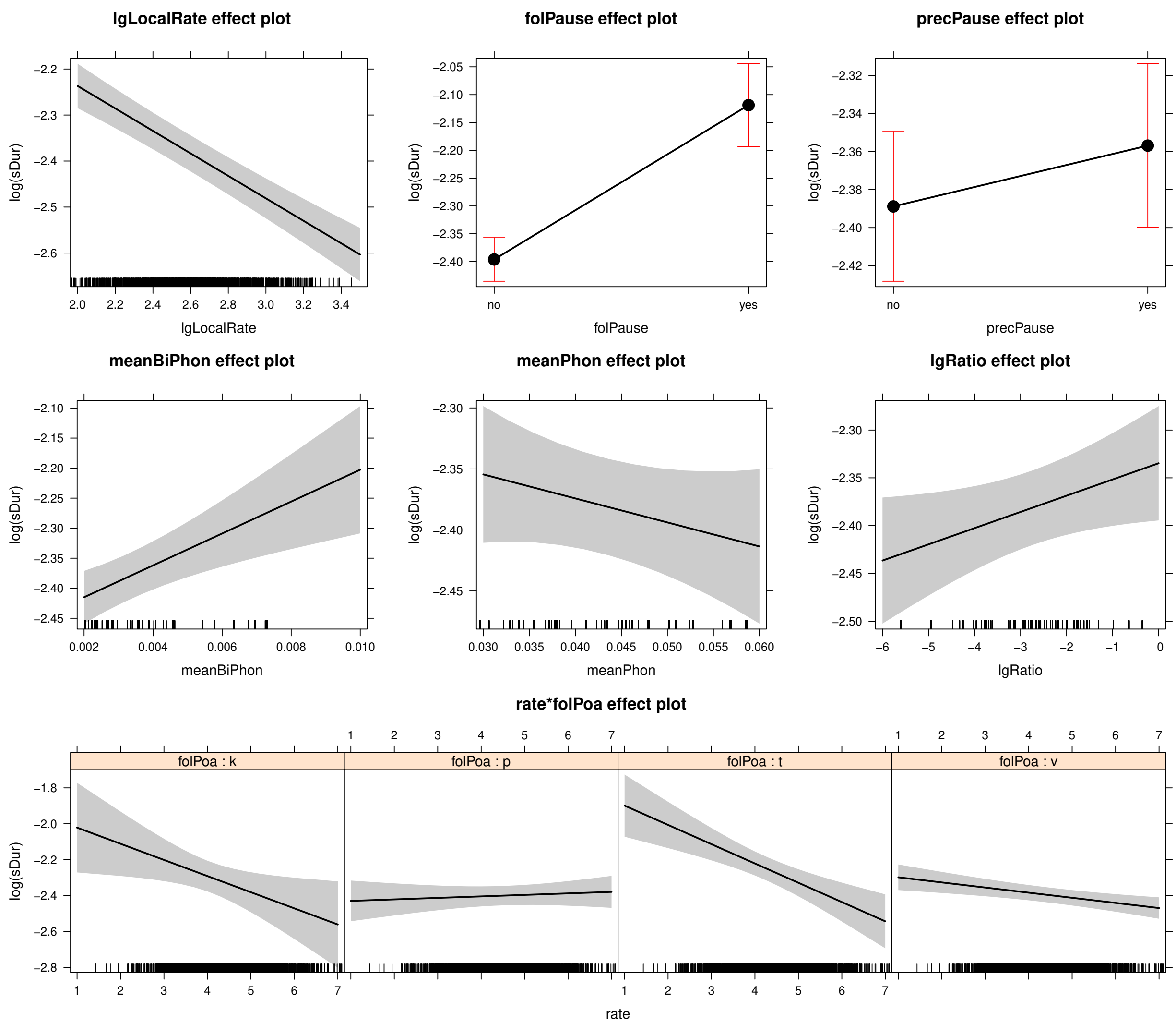


Figure 2.3: Partial effects plot of the suffix duration model. Increased relative frequency of the singular yields longer suffix duration (middle right panel).

The model allowed only one random slope before failing to converge. The term which yielded the largest drop in AIC was a slope for **folPause** by verb. Adding this random effect had no appreciable effect on the coefficient estimates for the key predictors. After the addition of this random slope, all observations with scaled residuals greater than 2.5 standard deviations from 0 were removed (Baayen, 2008), and the model was refit. The data trimming resulted in the loss of 35 observations, or 1.7% of the data, leaving 1963 observations. The final model is given in the bottom row of Table 2.4. This adjustment did not change the coefficient size or improvement in model fit associated with **lgRatio** over the control model in any appreciable way. It is this model that is illustrated in Figure 2.3.

The size of the effect of **lgRatio** on the suffix as calculated by the model is roughly 10 ms.

This can be seen by calculating the predicted log duration of the suffix when **lgRatio** is at its maximum and minimum values, and all other predictors in the model are held constant. For the current demonstration, all numerical predictors will be set at their mean values in the trimmed data set, and all factors will be set at their default values. This means that we are calculating the model's predicted log-transformed suffix duration for a word with a following alveolar stop and no preceding or following pause, with average phonotactic probability, spoken at an average rate. If such a word has the minimum value of **lgRatio** — in other words, it is extremely paradigmatically improbable — its predicted duration is -2.4 log units, or, exponentiated, 90.7 ms. If our hypothetical word has the maximum value of **lgRatio**, then its predicted duration is -2.29 log units, or, exponentiated, 101.2 ms. These values nicely bracket the actual mean suffix duration in the trimmed data set, which is 96.0 ms, and the difference between them, 10.46 ms, is the effect size of relative frequency on suffix duration.

2.3.2.2 Model evaluation

Figure 2.4 displays two plots demonstrating the behavior of the final model residuals. In the left pane, the actual suffix durations are plotted against the durations predicted by the model. The roughly constant size of the point spread indicates that the model fit is relatively homogeneous: The accuracy does not change across different values of the predicted value. The right hand pane shows the a qq-plot of the scaled residuals, and the close adherence of the qq-plot to the line $y = x$ indicates that the residuals are normally distributed, which a Shapiro-Wilk test for normality confirms: $W = 0.9986, p > 0.1$ (Baayen, 2008).

A further indication of model fit is provided by the variants of the statistic R^2, which indicates the proportion of variance explained by the model. Nakagawa and Schielzeth (2013) propose two types of R^2, designed to apply specifically to mixed-effects models. The conditional R^2 represents the total proportion of variance explained by both the fixed effects and the random effects of the model. For the refit model containing only **lgRatio** in addition to the control predictors and random slope, the conditional R^2 was 0.503. In other words, just over half of the total variance in the data is explain by this model. Of more interest in the present case, however, is Nakagawa and Schielzeth (2013)'s marginal R^2, which represents the variance explained solely by the fixed effects. This figure is given for each model in the rightmost column of Table 2.4, along with the increase in R^2 associated with the addition of each predictor. As can be seen in the rows associated with **ratioResid** in each of the four lower blocks (9a, 10b, 10c, refit 9a+slope), adding relative frequency to the model explains about 0.2% of the variance in the model — hardly a large effect, but consistent across different models.

2.3.3 Suffix center of gravity

2.3.3.1 Model summary

The spectral center of gravity (CoG) ofthe *-s* suffix was analyzed with a mixed effects regression model built according to the procedure outlined in Section 2.2.5.3. The control predictors that

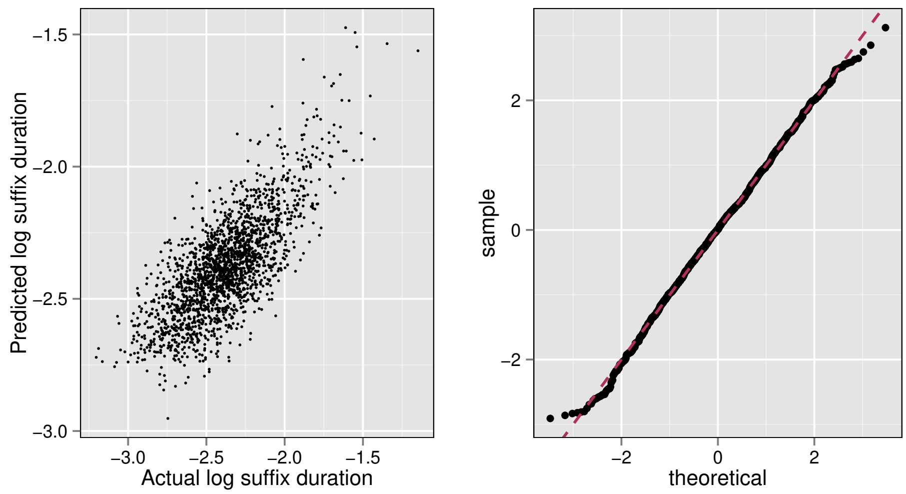


Figure 2.4: Model evaluation plots of the suffix duration model. The left panel illustrates the homogeneity of residuals and the right plot shows normality of residuals.

improved model fit were **rate**, **folPause**, **folMoa**, and log-transformed suffix duration (**lgSDur**). These predictors had similar effects on center of gravity as they did on duration. Faster speaking rate yielded lower CoG, while a following pause raised it. The effect of the manner of articulation of the following consonant revealed a broad divide between vowels and stops: Before following vowels, CoG was higher than before following stops. Finally, longer suffix duration yielded higher CoG. No other control predictors significantly improved the model fit. These findings are detailed in the top block of Table 2.5, and illustrated in Figure 2.5

The key predictors were tested in the same way as in the model of suffix duration. No key predictor improved the fit sufficiently to justify the added complexity in the model. Their effects are summarized in the bottom three blocks of Table 2.5. Although the simpler fixed effects structure permitted a more elaborate set of random slopes, their addition did not change the apparent absence of effect for any of the key predictors, and so will not be discussed further here.

For each model containing an added key predictor (rows 5a-c in Table 2.5), the data were trimmed parallel to their treatment with the model for suffix duration, in order to determine whether the poor contribution of the key predictors was due to the confounding influence of a few outlier observations that skewed the model fit. For each key predictor, this resulted in the loss of 45 observations, or 2.2% of the initial set of 2001 observations. For each key predictor, the model was refit without the outlier data points. As with the suffix duration model, however, this did not result in any appreciable change in the effect of any of the key predictors. The final model, therefore, was simply the control model, refit after removing observations with outlier residuals.

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p	R^2 (ΔR^2)
0	**Intercept**	7606	311	24.46	32814			0
1	**rate**	-70	28	-2.50	32810 (-4)	5.64	< 0.05	0.002 (0.002)
2	**folPause**				32805 (-5)	6.77	< 0.01	0.002 (0.001)
	yes	140	89	1.58				
3	**folMoa**				32773 (-32)	34.59	< 0.001	0.015 (0.012)
	v	416	64	6.49				
4	**lgSDur**	268	86	3.13	32765 (-8)	9.79	< 0.01	0.019 (0.004)
5a	**lgRatio**	-28	28	-0.99	32766 (+1)	1.00	0.32	0.019 (0.001)
5b	**entropy**	-141	111	-1.28	32765 (0)	1.66	0.20	0.019 (0.000)
5c	**sap**	9	11	0.78	32677 (+2)	0.60	0.44	0.019 (0.001)

Table 2.5: Center of Gravity model of the sufix. Blocks 5a-c indicate the effect of adding each key predictor to the control model (line 4). ΔR^2 was calculated before rounding.

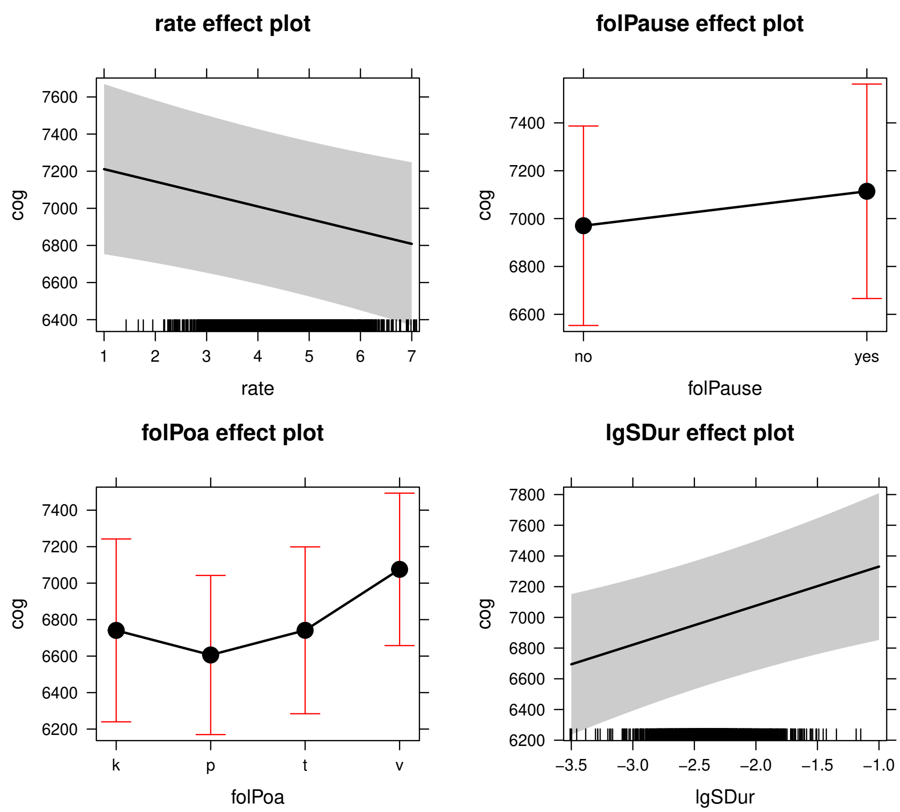


Figure 2.5: Partial effects plot of the control model shown in line 4 of Table 2.5.

2.3.3.2 Model evaluation

Despite the poor performance of the fixed effects in this model, the fit even of the control model without any random slopes (row 4 in Table 2.5) was quite tight. Parallel to Figure 2.4, the panels in Figure 2.6 show the observed CoG plotted against the predicted CoG from the control model on the left, and a qq-plot of the residuals on the right. Despite the tightness of the model fit, the residuals are not quite as well-behaved: There is a bit more scatter towards the higher CoG values, and a Shapiro-Wilk test shows that the model residuals are not normally distributed ($W = 0.996, p < 0.001$).

A comparison of the marginal and conditional R^2 values of the control model (Nakagawa and Schielzeth, 2013) reveals that almost all of the explained variance is due to the random intercepts. The marginal R^2 of 0.019 for the control model shows that fixed effects are explaining less than 2% of the variance, while the conditional R^2 shoots up to 0.72 by including variance explained by the random intercepts. In the refit control model, the imbalance is even stronger, with a marginal R^2 of 0.016, and a conditional R^2 of 0.79. In fact, in a null model containing only a random intercept for subject, the conditional R^2 is already at 0.69. Evidently, the majority of variation in spectral center of gravity is due primarily to individual differences in pronunciation, rather than contextual variation.

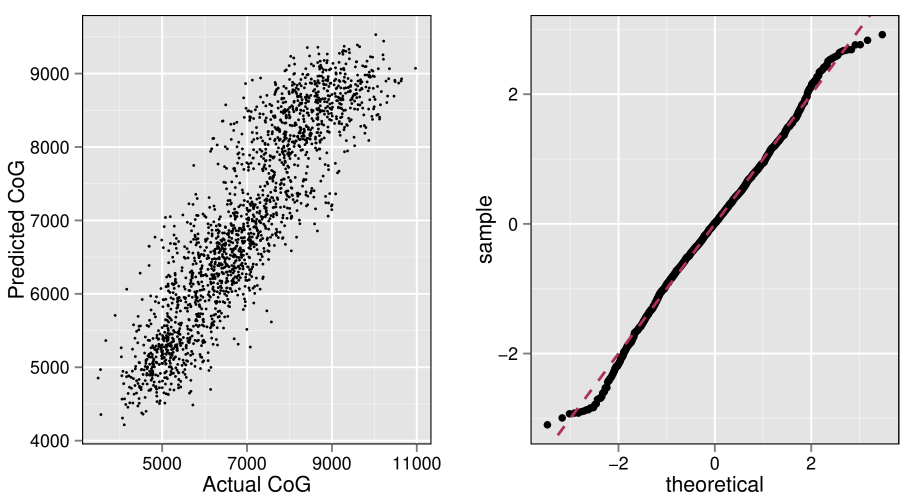


Figure 2.6: Model evaluation plots of the suffix center of gravity model. The left panel illustrates the homogeneity of residuals and the right plot displays a qq-plot of the residuals.

2.3.4 Stem duration

2.3.4.1 Model summary

The control model of log verb stem duration was analyzed with a mixed effects linear regression model built according to the procedure outlined in Section 2.2.5.3. Fixed effects were **nPhon**,

rate, **folPause**, **precPause**, **folMoa**, **meanBiPhon**, **meanPhon**, and **lgCumFq**. Many effects on the stem duration were similar to their effects on suffix duration. As speaking rate increased, duration decreased. The presence of a following pause yielded substantial lengthening, while the presence of a preceding pause yielded moderate shortening. Following vowels resulted in a slightly shorter verb stem, and higher phonotactic probability — both monophone and diphone — yielded insignificant lengthening. During the initial stages of the model building procedure, the addition of **meanBiPhon** and **meanPhon** improved model fit, but the magnitude of their estimated coefficients shrank after more predictors were added. Finally, as cumulative verb frequency increased, stem duration shrank. These effects are detailed in the top block of Table 2.6, and illustrated in Figure 2.7.

The key predictors, **lgRatio**, **entropy**, and **sap** were each tested to see how they improved the fit of the control model. As with the suffix duration model, the addition of **entropy** did not contribute anything to model fit, whether it was added directly to the control model (row 9b in Table 2.6), or residualized and added to the model already containing **lgRatio** and **sap** (row 11a). By contrast, an increase in **lgRatio** resulted in a shorter stem duration ($\beta = -0.03, SE(\beta) = 0.01, t = -2.15$), which did improve model fit ($\chi^2 = 4.76, p < 0.05$). This effect held regardless of whether it was added directly to the control model (row 9a) or residualized and added to a model already containing **entropy** (row 10b). By itself, **sap** did not improve model fit ($\chi^2 = 0.99, p = 0.32$), but in interaction with cumulative verb frequency (**lgCumFq**), it yielded a powerful improvement: ($\chi^2 = 34.79, p < 0.001$). As can be seen in the lower pane of Figure 2.7, at lower cumulative frequencies, higher **sap** resulted in shorter verb stem duration, while the effect was reversed for the highest frequency verbs. In other words, as singular agreement became less likely, the stems of low-frequency singular verbs became shorter, while the stems of high-frequency singular verbs became longer. Both **lgRatio** and the interaction of **sap** with **lgCumFq** remained significant when each was added to a model already containing the other (rows 10a and 10c in Table 2.6). This fixed effects structure — the control model plus **lgRatio** and the interaction of **sap** with **lgCumFq** — was retained for all further analyses.

The model containing both **lgRatio** and the interaction of **sap** and **lgCumFq** could accomodate two random slopes, both grouped by verb. These were **rate** and **folPause**. Adding these random slopes did not change the effects of either the control predictors or the key predictors in any appreciable way. Next, observations with scaled residuals more than 2.5 standard deviations from 0 were discarded, resulting in the loss of of 39 observations, or about 1.9% of the data. Finally, the model was fit to this trimmed data set one last time. There was still no appreciable change in the effects of either **lgRatio** or the interaction of **sap** with **lgCumFq**. Both significantly contributed to model fit, as shown in the final block of Table 2.6. It was this model whose partial effects are plotted in Figure 2.7. The details of the key predictors in this model are given in the bottom block of Table 2.6.

2.3.4.2 Model evaluation

Figure 2.8 shows the behavior of the residuals of the trimmed full model detailed in the bottom block of Table 2.6. The left panel plots the observed stem duration against the predicted stem

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p	R^2 (ΔR^2)
0	**Intercept**	-1.1	0.13	-8.50	-1200			0.00
1	**nPhon**	0.06	0.03	2.1	-1206 (-6)	7.20	< 0.01	0.033 (0.033)
2	**rate**	-0.08	0.01	-13.37	-1296 (-90)	91.64	< 0.001	0.082 (0.049)
3	**folPause**				-1358 (-62)	63.87	< 0.001	0.103 (0.022)
	yes	0.14	0.02	8.71				
4	**precPause**				-1458 (-100)	102.73	< 0.001	0.131 (0.028)
	yes	-0.11	0.01	-10.32				
5	**folMoa**				-1463 (-5)	7.13	< 0.01	0.137 (0.006)
	v	-0.05	0.02	-3.08				
6	**meanBiPhon**	6.60	11.98	0.55	-1465 (-2)	3.88	< 0.05	0.160 (0.022)
7	**meanPhon**	2.47	2.29	1.08	-1468 (-3)	4.47	< 0.05	0.179 (0.020)
8	**lgCumFq**	-0.04	0.01	-5.99	-1481 (-13)	15.46	< 0.001	0.250 (0.071)
9a	**lgRatio**	-0.03	0.01	-2.15	-1484 (-3)	4.76	< 0.05	0.270 (0.019)
10a	**sap**	-0.04	0.01	-5.87				
	sap*lgCumFq	0.005	0.001	5.88	-1516 (-32)	36.01	< 0.001	0.288 (0.018)
11a	**entropyResid**	0.07	0.06	01.03	-1515 (+1)	1.17	0.28	0.292 (0.004)
9b	**entropy**	-0.01	0.06	-0.15	-1480 (+4)	0.39	0.53	0.0251 (0.000)
10b	**ratioResid**	-0.05	0.02	-2.70	-1483 (-3)	5.57	< 0.05	0.274 (0.024)
9c	**sap**	-0.04	0.01	-5.87				
	sap*lgCumFq	0.005	0.001	5.88	-1513 (-32)	35.78	< 0.001	0.268 (0.018)
10c	**lgRatio**	-0.03	0.01	-2.15	-1516 (-3)	4.99	< 0.05	0.288 (0.020)
refit 8	**rate** and **folPause** by **verb**				-1899			0.176
refit 9c	**sap**	-0.04	0.01	-6.00				
	sap*lgCumFq	0.004	0.001	5.94	-1932 (-33)	36.90	< 0.001	0.196 (0.020)
refit 10c	**lgRatio**	-0.03	0.01	-2.43	-1936 (-4)	6.20	< 0.05	0.212 (0.016)

Table 2.6: Stem duration model summary. Blocks 9a - 10c show the effect of adding key predictors to the control model (row 8) in the order indicated. The final block shows the effect of adding key predictors to the control model with the random slopes **rate** and **folPause** grouped by verb, after trimming observations with outlier residuals. ΔR^2 was calculated before rounding.

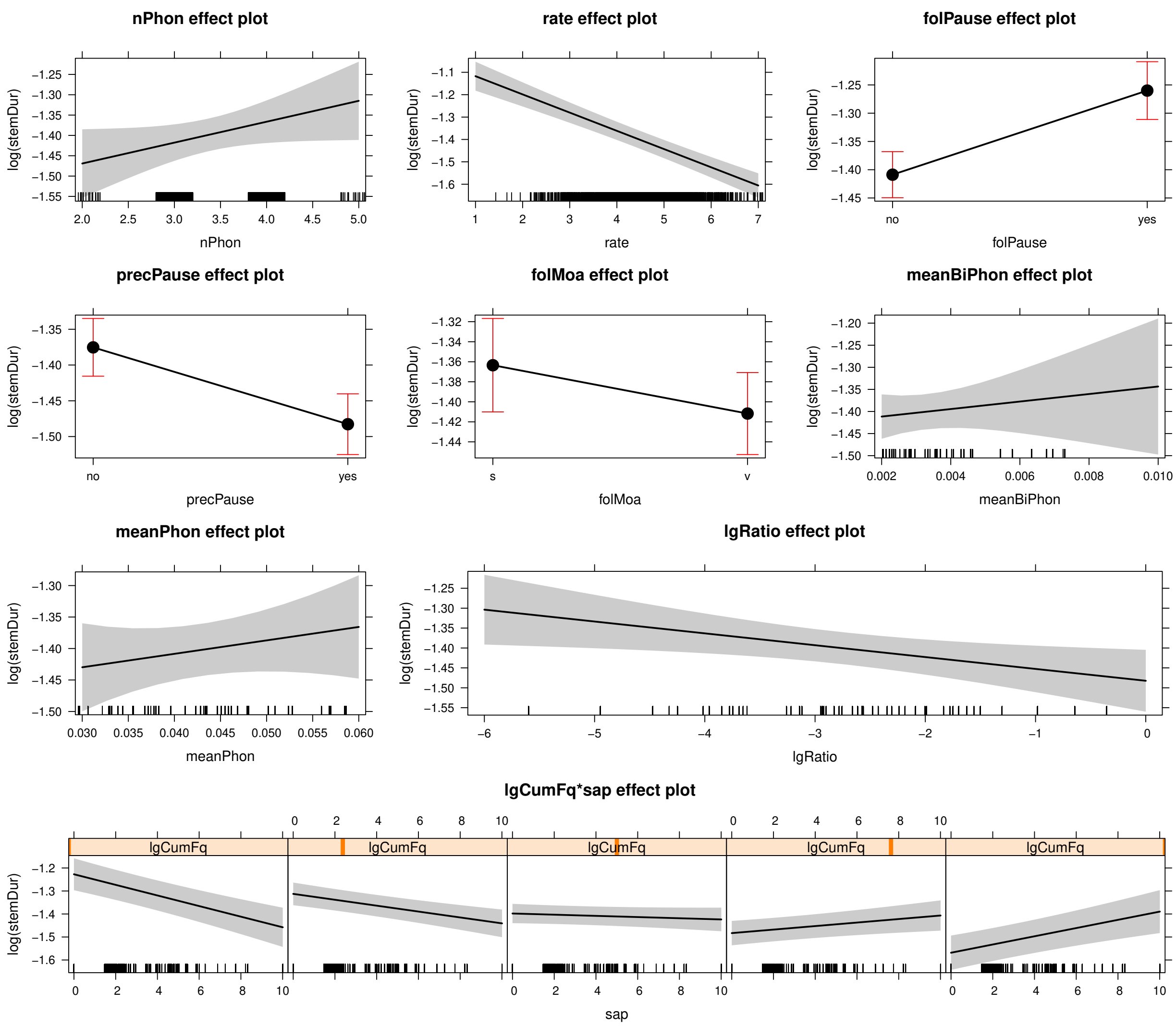


Figure 2.7: Partial effects plot of the stem duration model. Increased relative frequency of the singular yields shorter stem duration (right panel, third row), while probability of plural agreement interacts with cumulative verb frequency (bottom row).

duration. As with the other models, the plot indicates reasonable homogeneity: No interval shows more scatter in the predicted values than any other interval. The qq-plot in the right-hand panel of Figure 2.8 shows that the residuals are nearly, but not quite normally distributed, an observation confirmed by a Shapiro-Wilk test: W= 0.9982, $p < 0.05$.

The marginal R^2 of the full model (Nakagawa and Schielzeth, 2013) is 0.212, indicating that just over a fifth of the variance in stem duration is accounted for by the fixed effects. The key predictors together explain about 3.6% of the variance — 2% from the interaction of **sap** with **lgCumFq**, and 1.6% from **lgRatio**. This means that about 16% of the variance that can be explained by the fixed effects of this model is explained by the key predictors (0.034/0.212). The

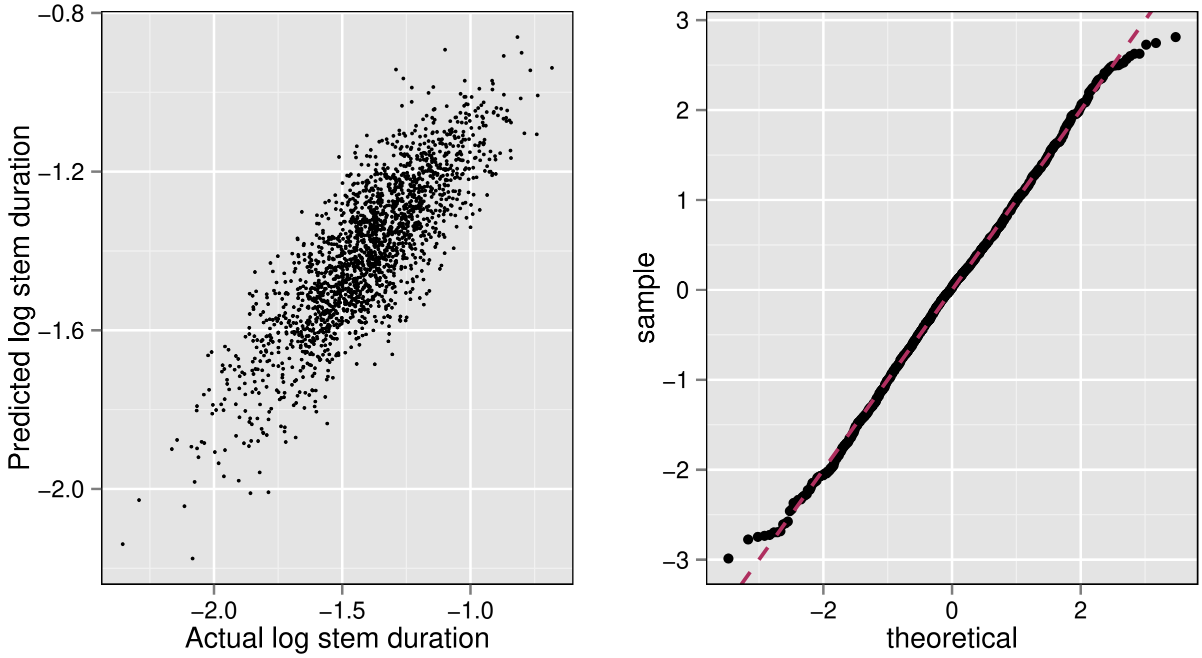


Figure 2.8: Model evaluation plots of the stem duration model. The left panel illustrates the homogeneity of the residuals, and the right plot shows near normality.

conditional R^2 is 0.766, indicating that fixed and random effects together explain over three quarters of the variance in stem duration.

2.4 Discussion

In a production experiment, pronunciation of singular present-tense verbs varied according to both contextual and paradigmatic probability. In particular, the pronunciation variation was confined to duration. Spectral center of gravity did not show sensitivity to either type of probability. The effect of contextual probability was confined to the verb stem: As singular agreement became more probable in the context of the sentence, singular verbs from low-frequency paradigms had longer verb stems, while those from high-frequency paradigms had shorter stems. There was no such effect on the suffix. The effect of paradigmatic probability was carried by the relative frequency of the singular form to the plural form of the verb, and appeared both on the suffix and the verb stem. Higher paradigmatic probability lengthened the suffix, and shortened the verb stem. There was no apparent effect of inflectional entropy.

On the basis of these findings, it is now possible to address the key questions that motivated this chapter.

2.4.1 Answers to key questions

Question 1 *Are contextually probable singular agreement suffixes phonetically reduced?*
No and yes. Contextual probability does not affect absolute suffix duration, but its effect on stem duration has consequences for the proportion of the entire duration of the verb form that is occupied by the suffix. The importance of relative final duration differences has been proposed already by Lunden (2010, 2013) as an explanation for apparent extrametricality effects of final consonants in English and Norwegian. In these languages (and many others), final CVC syllables behave as if they are lighter than medial or initial CVC syllables. They cannot attract stress, for example, while medial CVC syllables can. Lunden proposes that this phenomenon is a general consequence of word-final lengthening: Word final light syllables have a rhyme with a single vowel, which is longer than vowels in their non-final counterparts. The addition of a single final consonant therefore lengthens a word-final syllable proportionally less than it lengthens a shorter non-final syllable. To achieve the same relative increase in syllable duration word-finally, it is necessary to add two final consonants. It is for this reason that CVC syllables count as heavy word-medially, while word-finally heavy syllables must have the form CVCC. This same logic can be applied to the current findings. Probabilistic variation in pronunciation of a final *-s* will be more salient if the duration of the *-s* varies with respect to the verb stem.

The hypothesis that the key effect on the suffix was a change in proportional duration was confirmed by a post-hoc analysis in which a regression model was built to predict the log-transformed duration ratio of the suffix to the stem. In all respects the model-fitting procedure was identical to the procedure described in Section 2.2.5.3, and the effects of both relative frequency and contextual probability confirmed the hypothesis. Both **lgRatio** and the interaction of **sap** and **lgCumFq** significantly improved fit. Increased relative frequency resulted in longer relative suffix duration, as to be expected from its lengthening effect on absolute suffix duration (Section 2.3.2) and shortening effect on absolute stem duration (Section 2.3.4). Similarly, the interaction of contextual probability cumulative verb frequency showed the expected inversion from its effect on stem duration alone. Increasing probability of singular agreement led to shorter relative suffix duration for verbs from low-frequency paradigms, while those from high-frequency paradigms had lengthened relative duration. In other words, contextually probable singular suffixes are relatively reduced for low-frequency verbs, but not for high-frequency verbs.

Question 2 *Do relative frequency and inflectional entropy have distinct effects on pronunciation?*
Yes: Increased relative frequency consistently yields longer suffix durations — both absolute and relative — while inflectional entropy seems to have no effect. Despite frequent findings of the importance of entropy on the timecourse of lexical retrieval (Baayen et al., 2006; Bien et al., 2011; Milin et al., 2009; Moscoso del Prado Martín et al., 2004; Tabak et al., 2005), there is no apparent concomitant effect on pronunciation. This finding provides an interesting counter to Baayen et al. (2008), who observed that relative frequency affected comprehension of inflected word forms, while inflectional entropy came into play during production. They propose that mechanisms for comprehending inflected forms are aided by high relative frequency, while mechanisms for pro-

ducing them carry a cost for accessing lexemes from high entropy paradigms, regardless of the relative frequency of the target word. By contrast, the current study found that production processes — indexed by pronunciation — seem insensitive to entropy, regardless of its temporal cost during lexical retrieval, and highly sensitive to relative frequency.

One possible account for this distinction is the task. In single-word paradigms of picture naming or lexical decision like those used by Baayen and colleagues, participants must select the lexeme that identifies a particular picture, or evaluate whether a particular orthographic form represents a valid word form in the language. In both cases, high level processing is necessary. Picture naming requires participants to recognize a semantic concept and select the appropriate lemma and the lexeme that encodes it before any aspect of articulation comes into play. Lexical decision similarly involves semantic processing in a way that simple word-naming does not (Baayen et al., 2007). These types of tasks might therefore require participants to draw upon a large amount of stored lexical knowledge, which includes paradigmatic structure and inflectional entropy. Further, the single-word design insulates participant performance from other factors that might obscure the effects of interest. By contrast, in the current experiment the desired form of the verb was provided to participants well before they produced it. The fact that they did occasionally produce plural verbs even when provided with the singular verb suggests that they were computing agreement relations independently of the written stimulus. Yet the presence of the target verb may have aided them in navigating the paradigm structure of the target lexeme as they retrieved the necessary word form. The result of these task differences is a difference in which stages of lexical retrieval were most active. In single-word picture naming and lexical decision tasks, higher-level processing renders retrieval vulnerable to inhibitory effects of inflectional entropy. By contrast, in the current task, high-level conceptual processing was minimized, and any effects of entropy may have been further limited or obscured by many other simultaneous processes — such as agreement computation and the retrieval of surrounding words. At the same time, the analysis focused on pronunciation — the output of lower-level phonological encoding and articulation — and revealed the effects of relative frequency. Under this account, inflectional entropy and relative frequency do have distinct effects on lexical retrieval, but the distinction is not solely a difference between perception and production, as Baayen et al. (2008) propose. Rather, it also incorporates a distinction between high-level processing, at which entropy comes into play, and low-level processing, when relative frequency is more important.

2.4.2 The role of cumulative frequency

The cumulative frequency of all forms of the verb explained about 7.1% of the variance in the control model for verb stem duration, and an additional 1.8% when it interacted with contextual probability. Overall, the effect of cumulative frequency was reduction: More frequent forms had shorter stems, which translated into relatively longer suffixes. Yet it also modulated the effect of contextual probability. Whereas low frequency contextually probable *-s* was relatively shorter, compared to the stem, high-frequency contextually probable *-s* was longer. What does this interaction mean?

For low-frequency verbs, the association between increased contextual probability and reduced duration of the suffix is entirely expected. As laid out in Section 2.1.1.1, phonetic reduction is a frequent hallmark of contextually probable units. Here, the unit in question is the agreement suffix. In Sg-Sg sentences with singular head nouns and singular local nouns (e.g., *The pizza with the missing slice looks unappetizing* . . .), the contextually probable aspect of the verb is not the stem (*look*) or the entire word form (*looks*), but the suffix alone (-*s*). The fact that the probability targets the suffix alone means that the suffix will be most sensitive to probability when it has the most independence during retrieval. According to dual-route models of lexical retrieval, it is exactly the low-frequency complex forms that are most likely to be retrieved with some degree of decompositionality, especially when they are read (Baayen et al., 1997; Baayen and Schreuder, 1999). For middle-frequency verbs, whole-word representations have more of an advantage, which damages the independence of the *-s* suffix, and makes it less sensitive to the contextual probability. This is because it is not the whole word (*looks*) that is probable in Sg-Sg sentences, but the suffix alone. The result of this advantage for whole-word retrieval is a reduction in the effect of contextual probability on relative suffix duration. What is perplexing about this interaction is the reversal of the effect for high-frequency verbs. Why would a highly probable singular verb, retrieved via a whole-word representation, show a longer relative suffix duration?

One possibility is that the target of probabilistic pronunciation variation switches from the suffix, when the word is retrieved in a decomposed manner, to the entire word form, when the word is retrieved through a whole-word representation. In other words, the apparent interaction of frequency with contextual probability that appears on the stem (bottom panel of Figure 2.7) in fact represents two different processes. For low-frequency verbs (shown in the left-most two panes), the lengthening of the stem that comes with increased contextual probability is a way of decreasing the relative duration of the suffix — phonetic reduction. For high-frequency verbs (right-most two panes), the shortening of the stem is a way of decreasing the duration of the entire singular verb, which was retrieved whole in a context where singular agreement is highly probable — again, phonetic reduction.

The idea that any singular verb form can show reduction in contexts where singular agreement is probable has been observed before. Kuperman and Bresnan (2012), for example, found reduction in ditransitive verbs and their following nouns when they appeared in the more probable variant of the dative alternation. Similarly, Demberg et al. (2012) found reduction in words that appeared in more probable syntactic structures. In other words, it was not the specific unit whose probability led to reduction; it was the abstract syntactic frame in which the unit appeared. The current finding — that verbs which carry the most probable agreement features have shorter stems — is one more instance of this phenomenon.

If the change from lengthening the stem to shortening the stem represents the switch from decomposed retrieval to whole-word retrieval, then the crossover point appears to be at about the median value of cumulative verb frequency. This corresponds to roughly 14 words per million.[3]

[3]The median log frequency is 6.58, which, exponentiated, gives a frequency of 721 in the SUBTLEX corpus (Brysbaert et al., 2012). The corpus contains 51 million words, which means that the normalized median frequency is 721/51 = 14.14 words per million.

This threshold is more stringent than the 6 words per million proposed by Alegre and Gordon (1999) (although see Baayen et al., 2007, for an alternative view). Although 14 words per million is larger, it represents the cumulative frequency of the entire verbal paradigm. The frequencies of the singular verb forms from paradigms with cumulative frequencies of about 14 per million (*books*: 0.25, *bumps*: 0.41, *taps*: 0.76, *tips*: 1.57) are all much lower than 6 per million .

2.4.3 Relative frequency: enhancement or reduction?

The original intent of examining the effect of inflectional entropy separately from relative frequency was to see whether such a distinction could account for the fact that complex word forms with high relative frequencies compared to their stems have shown both phonetic reduction and phonetic enhancement. Hay (2003), recall, found that derived adverbs like *softly* and *swiftly* are more likely to have the [t] deleted when the adverb is more frequent with respect to the base. Thus, *swiftly*, which is quite frequent with respect to *swift*, is pronounced without the [t] more often than *softly*, which is less frequent compared to *soft*.[4] By contrast, Schuppler et al. (2012) found the opposite effect in past-tense Dutch verbs: As the relative frequency of the inflected form to the stem increased, the final *-t*, which marked past tense, was *less* likely to be deleted. The authors propose that one explanation for these differing results is the distinction between reduction of the stem and reduction of the suffix. The deleted *-t* in Hay (2003)'s experiment was part of the adjective stem, while the deleted *-t* in Schuppler et al. (2012)'s study was the suffix itself. Perhaps the phonetic enhancement associated with increased paradigmatic probability, also observed by Kuperman et al. (2007) and Hay et al. (2012), applies only to affixes. The current experiment supports this distinction. The relative frequency of the singular verb form to the plural verb form also happens to be the relative frequency of the inflected verb form to the stem, which renders this measure comparable to those used by Hay (2003) and Schuppler et al. (2012). As the relative frequency increased, the duration of the suffix increased — affix enhancement that is parallel to Schuppler et al. (2012). Yet, in addition, as the relative frequency increased the duration of the stem decreased — stem reduction that is parallel to Hay (2003)'s.

2.4.4 Lingering questions

A limitation of the current study arises from the fact that English has only one subject agreement morpheme: the singular *-s* observed here. Yet the current findings have been interpreted as describing pronunciation variation in agreement morphemes in general, not simply the singular agreement suffix. If this interpretation is accurate, then the behavior for the singular suffix should be mirrored in the behavior for the corresponding plural suffix. English has no distinct plural suffix — a property that might in fact be responsible for the effects of contextual probability. It is therefore crucial

[4]Hay (2003) uses word frequency measurements from the CELEX lexical database (Baayen et al., 1993), in which *swiftly* is more frequent than *swift*. According to data from the SUBTLEX corpus, however, both adverbs are less frequent than their bases, but it is nevertheless true that the relative frequency of *swiftly* to *swift*, 83/155 = 0.53, is higher than the relative frequency of *softly* to *soft*, 240/1631=0.15.

to determine whether these effects translate into other languages that do have distinct singular and plural agreement suffixes.

A second question arises from the apparent null effect of inflectional entropy on suffix pronunciation. As proposed earlier, it could indeed be the case that inflectional entropy affects speech onset latency (Bien et al., 2011), but not pronunciation, due to differences in task demands. Yet it could also be the case that the verbal inflectional paradigm in English is so simple that it does not allow inflectional entropy a wide enough range of variation for an effect to emerge. This is the account that Bien et al. (2011) propose to explain why they observed an inhibitory effect of inflectional entropy in a production task involving Dutch verbs, but not Dutch adjectives. Where Dutch adjectives have only two possible forms — singular and plural — the verbs have a much richer paradigm. Similarly, where English verbs allow only four (or, at most, five) forms, a language with more verbal forms might have a sufficient range of inflectional entropy in its verbal paradigms for an effect to emerge in verbal pronunciation.

In order to address these two questions, it is necessary to explore a language with the following two characteristics: There must be distinct suffixes marking both singular and plural agreement, and the verbal inflectional paradigm must be substantially richer than English's. Russian meets both of these requirements, and the next two chapters will therefore examine agreement production in Russian. Chapter 3 is a norming study, in which I build the computational tools necessary to calculate the contextual probability of a singular or plural suffix in the sentence, and Chapter 4 presents the results of a production experiment parallel to the experiment presented here.

2.5 Conclusion

This chapter revealed that the pronunciation of singular *-s* suffixes in English verbs are sensitive to different types of probabilities. When they encode more probable agreement relations in the sentence, they are shorter, relative to the duration of the stem, or the stem itself is shorter. By contrast, the absolute duration of the suffix is lengthened when they are more frequent within their inflectional paradigms. These findings both corroborate previous research into the effects of contextual probability and paradigmatic probability, and expand them, by showing that pronunciation of individual morphemes varies in the same way segments, syllables, and whole words do. The interaction between verb frequency and both types of probability further highlights the complexity of word form retrieval during speech production. The retrieval of words as whole forms or as component parts is dependent on how frequently the word is used. It is a testament to human cognitive capacity that we are able to track these patterns during speech as quickly as fluently as we do.

Chapter 3

Agreement variation in Russian: A multivariate analysis

3.1 Introduction

Language is full of grammatical constructions which permit a certain degree of variability in their expression. To produce even the most mundane utterances, language users must navigate this variability to communicate their thoughts. When asking about the contents of the pantry, for example, a speaker might say *Who forgot to put away the crackers*, even though it would have been equally acceptable to say *Who forgot to put the crackers away?* The addressee might respond *I gave the dog the leftover crackers*, an utterance that rejects the alternative option of *I gave the leftover crackers to the dog.* How do speakers choose which grammatical realization to use?

Previous research in grammatical variation — especially in English — has determined that such choices are usually influenced by multiple variables. Rates of copula contraction and deletion in African American Vernacular English, for example, have long been known to be sensitive to the phonological makeup and pronominality of the subject, as well as the syntactic category of the following phrase (Baugh, 1980; Labov, 1969), and verb particle placement is affected by the length of the direct object, the type of determiner it has, and the idiomaticity of the VP, among many others (Gries, 2003). A similarly large constellation of factors has been shown to be relevant in predicting whether a verb will be used in active or passive voice (Weiner and Labov, 1983), which relative pronoun will be used (Guy and Bayley, 1995), which form of the dative alternation will be used (Bresnan et al., 2007), whether a possessive construction will employ the *-s*-genitive or the *of* genitive (Szmrecsanyi and Hinrichs, 2008), and when a subject personal pronoun will be used or omitted in Spanish (Erker and Guy, 2012).

One type of grammatical variation that is already of enormous interest in psycholinguistic research is subject-verb number agreement, which has been the subject of study in many languages, including English (Eberhard et al., 2005), Dutch (Antón-Méndez and Hartsuiker, 2010), French (Franck et al., 2002), Hebrew (Deutsch and Dank, 2009), Basque (Santesteban et al., 2013), Serbian (Mirković and MacDonald, 2013), and Russian (Lorimor et al., 2008). This sort of variation is also sensitive to multiple variables: The choice between singular and plural copular forms in English existential constructions, for example, is influenced by the type of modifier on the subject, verb tense, and how long the sentence is after the copula (Hay and Schreier, 2004; Riordan, 2007). The goal of the current chapter is to extend this approach to another source of agreement variation: Russian quantified subject noun phrases.

Russian is an inflectionally complex language, with a strong system of grammatical agreement that pervades most constructions. Usually, the rules governing agreement relations are straightforward. Adjectives and determiners obligatorily agree with nouns in gender, number, and case, and verbs obligatorily agree with subjects in person and number (non-past tense) or in number and gender (past tense). Yet in some constructions, these rules become more complex. In (15-16), the identical subject noun phrase appears in the same text with two different agreement suffixes on the verb: Neuter singular *-o* or plural *-i* . (Simonov, 1958, cited and translated in Robblee, 1993a).

(15) U nego **sgore-l-o** dva tanka i v odnom iz nix — ves'
By him.GEN **burn-PAST-N.SG** two tank.GEN and in one.LOC from them.GEN — whole.NOM
ekipaž.
crew.NOM
'Two of his tanks had burned, and in one of them, the whole crew.'

(16) On naxodils'a na tol'ko čto vz'atom barxane, gde segodn'a **sgore-l-i**
It be.located.PAST.M.SG on just.now taken.LOC dune.LOC, where today **burn-PAST-PL**
dva tanka Klimoviča.
two tank.GEN Klimoviç.GEN
'It was located on the dune that had just been taken, where two of Klimovič's tanks had burned.'

Sentences with quantified subject noun phrases, such as *two tanks* or *several boys*, can occur with either singular or plural agreement, depending on a number of properties of the surrounding utterance. In the sentences above, for example, the quantifier *two* favors plural agreement in (16), while singular agreement is favored by the fact that the subject, *two tanks*, is inanimate and precedes the verb in (15) (Timberlake, 2004). The full set of these factors is quite wide-ranging, encompassing such variables as the grammatical category of the predicate (Corbett, 1979, 1998; Timberlake, 2004), the meaning of the predicate (Robblee, 1993a, 1997; Timberlake, 2004), the quantifier (Patton, 1969; Suprun, 1969), the word order (Corbett, 1983, 2006), the information structure of the utterance (Lambrecht and Polinsky, 1997; Nichols et al., 1980), the tense of the verb (Kuvšinskaja, 2012), the presence of other modifiers besides the quantifier (Corbett, 1979; Timberlake, 2004), the form of the other modifiers (Corbett, 2006; Suprun, 1957), and many, many others (Kuvšinskaja, 2012; Patton, 1969).

Existing studies of this particular phenomenon have tended to focus on the effects of one variable at a time, an approach that has limitations on two fronts. First, single-variable analyses usually focus on the behavior of a single variable by combining the observations across all possible values for the others. As a result, this approach disregards the fact that often certain variables tend to co-occur, making it difficult to determine whether the factor under consideration is actually responsible for the observed pattern. For example, Robblee (1993a) identifies three semantic classes of verbs that differ depending on the amount of agentivity of the subject. This study revealed that verbs with more agentive meanings have the highest rates of plural agreement. Yet it is very commonly observed that animate subjects also tend to have higher rates of plural agreement (Corbett, 2006; Kuvšinskaja, 2012; Patton, 1969; Timberlake, 2004). In fact, Robblee (1993a) remarks that more agentive verbs have a much more restricted set of plausible subjects — specifically, animate subjects. Is the effect of verb semantics, then, carried by the fact that the most agentive verbs tend to be used with animate subjects? Or could it be that the effect of animacy is carried by the fact that animate subjects tend to be used with agentive verbs? Since the discussion in Robblee (1993a) does not provide a breakdown of subject animacy across the different verb categories, it remains unclear whether the effects of animacy and verb agentivity are truly independent of each other.

Patton (1969) encounters a similar situation when discussing the agreement patterns of time expressions. Subjects which denote periods of time prefer singular agreement in literary prose (21% plural agreement rate), but then they are used with predicate-subject word order about 82% of the time. Since predicate-subject order itself prefers singular agreement, these proportions do not reveal whether it is the word order or the semantics of the subject that increase the preference for singular agreement, or whether they both have independent effects.

A more subtle example of the problems arising from a one-variable approach can be found in Kuvšinskaja (2012)'s observation that the presence of delimiting adverbs such as *only, precisely, almost* decreases the likelihood of plural agreement. Whereas the corpus in this study had an overall plural agreement rate of 71% (out of 1063 sentences with numerically quantified subjects), the set of 131 sentences containing such adverbs had a plural agreement rate of only 51%. There is no information provided, however, about the make-up of the set of sentences containing those adverbs: Do those 67 sentences with singular agreement share a higher rate of some other factor that could be influencing the agreement? For example, it is possible that sentences with lexical items delimiting the specific number of the subject are sentences whose primary informational purpose is to introduce that subject into the discourse, with particular focus on the specific quantity of the subject. This particular presentational semantics is associated with verb-subject word order (Timberlake, 2004), and verb-subject word order is consistently identified as a property that reduces plural agreement (Corbett, 1983, 2006). In fact, in all the sentences provided as examples with delimiting adverbs in Kuvšinskaja (2012), the verb did precede the subject. Perhaps, then, the reduced rate of plural agreement should be attributed to the different information structure and word order, rather than to the presence of the delimiting adverbs. In other words, the appearance of one factor could itself have been triggered by some other property that conditioned the agreement choice (see Kuvšinskaja, 2012, for further examples).

The second, related limitation of single-variable approaches is the limited consideration of possible interactions between the different factors. For example, Patton (1969) found that animacy interacted with register, such that quantified animal subjects pattern with humans in literary prose, preferring plural agreement, while in journalistic prose they pattern with inanimates, and are more likely to take singular agreement. Similarly, Robblee (1997) found that the effect of subject-verb and verb-subject order was not consistent across all quantifiers. These findings, however, are still based on simple percentages of plural agreement rates, and only a few pairs of predictors have been examined for possible interactions. A more nuanced statistical investigation could confirm or disconfirm that the effect of one factor is the same in all sentences across different values of other factors.

It is now standard for studies of other types of agreement variation to embrace an analysis that takes into account the combined effects of multiple variables. Hay and Schreier (2004), for example, examine the historical development of subject-verb non-concord in 19th-century New Zealand English as a function of speaker gender, speaker birth date, verb tense, and subject type. Similarly, Riordan (2007) explored rates of non-concord in American English existential constructions as a function of the subject determiner type, sentence polarity, the presence of a plural *-s* on the subject, the presence of any disfluencies, the length of the sequence following the subject, the age and gender of the speaker, and the type of discourse. The current study was therefore designed with two goals in mind. The first is to better understand the joint effects on Russian agreement variation of multiple variables that have previously been examined seperately. To that end, the current experiment explored verbal agreement patterns with quantified subject noun phrases in Russian as a function of five different properties: The quantifier, the animacy of the subject, the semantic individuation of the verb, the word order of the sentence, and the tense of the verb.

The second goal of this project was to complement the existing research — all of which is based on corpus data — by probing more deeply those issues of Russian agreement variation that corpus research cannot itself resolve. The primary difficulty in this case is the fact that some combinations of factors implicated in the agreement choice are simply too infrequent to permit an adequate analysis. For example, agentive verbs are so frequently used with animate subject NPs that there are not enough naturally occuring cases of agentive verbs and inanimate subject to disentangle the effects of animacy from the effects of agentivity. A second, deeper issue concerns the extent to which usage data is representative of how people process language. Divjak (2008), for example, found that patterns of grammaticality judgements elicited in experiments do not always line up with usage patterns found in corpora, suggesting that a full understanding of how speakers process certain constructions cannot be found through usage data alone. For these reasons, a more complete picture of the issue at hand is found not simply in corpora, but in the combination of corpus and experimental research (Gilquin and Gries, 2009). Therefore, the data examined here come not from a corpus, but from an experiment in which Russian speakers completed a fill-in-the-blank task for sentences that balanced all two-way combinations of four of the five factors at issue: Quantifier, verb semantics, word order, and subject animacy. Where previous research has identified these factors largely through the use of corpus data, this project evaluates their effects on the experimental side.

3.2 Methods

3.2.1 Experimental conditions

The stimuli were designed to test the independence and interaction of five possible different predictors that were most consistently identified in the literature: Quantifier identity (Quantifier), animacy of the subject noun phrase (Animacy), word order (Order), semantic class of verb (Verb), and the tense of the verb (Tense). Three categories of quantifier were of interest: paucals, which consist of the numerals *two, three, four*, low general numerals (*five, ten, twenty*, etc.), and approximate quantifiers (*several, few, many*, etc.) (Timberlake, 2004). The specific Quantifiers used here included the paucal *dva* 'two,' the low general numerals *pjat'* 'five' and *des'at'* 'ten,' and the three approximate quantifiers *malo* 'few,' *mnogo* 'many' and *neskol'ko* 'several.' The possible Animacy values were Animate or Inanimate, and all noun phrases were carefully selected to avoid other sources of variation, as follows. To avoid the possible confound of whether animals, which are grammatically animate, pattern with humans or inanimates in experimental settings (Patton, 1969), all animate subjects were human. To avoid any possible confusion of gender realization on past tense verbs, all subject nouns were grammatically masculine. The possible values for Order were Subject-Verb (SV) or Verb-Subject (VS), and sentences consisted of three constituents: a subject, a verb, and either a direct object or a prepositional phrase, depending on the transitivity of the verb.[1] In SV sentences the verb immediately followed the subject, while in VS sentences the verb immediately followed the non-subject constituent. This positioned all verbs in sentence-medial position, which allowed the effect of word order to be studied without the possible confound of verb-position.

Finally, the semantic class of the verb was manipulated according to the categories proposed by Robblee (1993a). It is already common to distinguish two types of predicates: Activity or "dynamic" predicates, and predicates denoting presence or existence (Corbett, 1983; Kuvšinskaja, 2012; Patton, 1969). Robblee (1993a), however, divides predicates more finely, into three semantic classes: Inversion, Intransitive, and Agentive.[2] According to this division, Inversion predicates describe simple existence and location (e.g., *be*, *appear*, *be necessary*), with no agentivity or manner information; Intransitive predicates describe movement and posture, but their meanings include no volitionality (e.g. *grow, stand, lie*); and Agentive predicates describe activities that are carried out intentionally by the actor (e.g., *hit, write, participate*). These three categories correspond not only to increasing degrees of plural agreement, but also to the occurence of genitive subjects in negated sentences (Robblee, 1993b). The predictive power of these verb classes in a domain of grammatical variation entirely separate from subject-verb agreement suggests that this division of verb semantics is valid, and its three-way distinction provides a more fine-grained classification system than the more widely used two-way distinction.

A summary of the different values for each condition is given in Table 3.1. If the patterns

[1] In some cases both a direct object and a prepositional phrase or some other adverbial modifier were present, in order to improve the plausibility of the sentence.

[2] Actually, Robblee subdivides each of the main classes into two subclasses, yielding six categories of verb semantics, but for the sake of simplicity these subclasses have been collapsed in this study.

observed in the previous studies hold here, plural agreement rates should be higher for conditions that are lower in each column.

Quantifier	Verb	Animacy	Order	
malo 'few'	**Inversion** e.g. *naxodits'a* 'be located'	**Inanimate** e.g. *gruzovik* 'truck'	**Verb subject (VS)** e.g. 'In the driveway [stood]$_V$ [many trucks]$_S$.'	Higher rates of plural agreement ↓
mnogo 'many'	**Intransitive** e.g. *stojat'* 'stand'	**Animate** e.g. *student* 'student'	**Subject verb (SV)** e.g. '[Few students]$_S$ [were located]$_V$ by the board.'	
neskol'ko 'several'	**Agentive** e.g. *udarit'* 'hit'			
des'at' 'ten'				
pjat' 'five'				
dva 'two'				

Table 3.1: Values for each predictor used in designing the test sentences, with examples of items that would be classified under each value.

3.2.2 Materials and design

Because of the large number of conditions, a full factorial design crossing all values of Quantifier, Verb, Animacy and Order ($6 \times 3 \times 2 \times 2$) would have required each participant to respond to 72 critical sentences. It seemed unlikely that participants could avoid noticing such a large number of quantified subject noun phrases unless the experiment included a prohibitive number of filler stimuli, so the set of combinations included in this study was reduced by half, yielding a final set of 36 conditions. A list of the conditions included and the conditions omitted is given in Table 3.2.

To construct the experimental items, six verbs of each Verb type were selected, for a total of 18. The verbs in each condition did not differ significantly in log frequency ($F(2, 15) < 1, p > 0.7$), as determined by data from the Russian National Corpus (www.ruscorpora.ru). Each verb was used to construct two separate sentences — one with an animate subject, and one with an inanimate subject. The resulting 36 sentences were rotated through the 36 conditions, creating six sets of experimental stimuli. Table 3.3 illustrates how this rotation through the stimuli conditions is done, and a sample stimulus set is given in Appendix B.

In addition to these 36 sentences, 12 structurally similar sentences were included. These additional sentences all contained end-stressed verbs, selected solely on phonological grounds, rather than semantic grounds. As will be seen in Chapter 4, the result of this study is crucial in designing

Quantifier	Animacy	Order					
		SV			VS		
		Inversion	Verb Intransitive	Agentive	Inversion	Verb Intransitive	Agentive
‘two’	Animate	✓	X	✓	X	✓	X
	Inanimate	X	✓	X	✓	X	✓
‘five’	Animate	X	✓	X	✓	X	✓
	Inanimate	✓	X	✓	X	✓	X
‘ten’	Animate	✓	X	✓	X	✓	X
	Inanimate	X	✓	X	✓	X	✓
‘several’	Animate	X	✓	X	✓	X	✓
	Inanimate	✓	X	✓	X	✓	X
‘few’	Animate	✓	X	✓	X	✓	X
	Inanimate	X	✓	X	✓	X	✓
‘many’	Animate	X	✓	X	✓	X	✓
	Inanimate	✓	X	✓	X	✓	X

Table 3.2: Distribution of factors in experimental design. X-marks indicate that a particular condition was not included in the experiment. Sample sentences for each of the three verb types are given, to demonstrate how they changed across conditions.

a follow-up production experiment that parallels the experiment in Chapter 2, and one requirement of that follow-up experiment is the use of end-stressed verbs. The purpose of these verbs' inclusion here, therefore, was to determine whether end-stress had any unexpected effect on the pattern of number agreement. As it turned out, it did not, and so those additional stimuli will not be discussed further. However, since they contained quantified subjects and were extremely similar in structure to the critical sentences, they were treated as if they were critical sentences in the stimulus-list design.

The 36 critical sentences (or 48, including the structurally similar additional sentences) in each set were randomly ordered and interspersed with 96 fillers, such that no two test sentences were presented adjacent to each other, and no two test sentences were separated by more than three fillers. In order to distract participants' attention away from the structural properties of the sentences, each sentence completion task was followed with a word-association task. In this way, every critical sentence was separated from the preceding critical sentence by at least three other stimuli (a word association, a filler, and a second word association) and as many as seven (three fillers and four word associations).

Each test sentence contained a blank in the place of the verb. Immediately after the blank the intended verb was provided in the infinitive, which is also the citation form. Since conjugated Russian verb forms are never homophonous or homographic with infinitive forms, this method of

Order	Quant.	Anim.	Verb: Inv.	Verb: Int.	Verb: Ag.
SV	‘two’	Anim	Два журналиста ___ (быть) в классе ‘Two journalists ___ (be) in the classroom.’	X	Два партизана ___ (пробиваться) к своим ‘Two partisans ___ (get through to) their own people.’
		Inan	X	Два рисунка ___ (лежать) в коробке ‘Two drawings ___ (lie) in the box.’	X
	‘five’	Anim	X	пять зоологов ___ (лежать) на траве ‘Five zoologists ___ (lie) on the grass.’	X
		Inan	Пять мобильников ___ (быть) на полке ‘Five mobile phones ___ (be) on the shelf.’	X	Пять потоков ___ (пробиваться) сквозь щели ‘Five automatic styluses ___ (write) in red ink.’
VS	‘two’	Anim	X	На траве ___ (лежать) два зоолога ‘On the grass ___ (lie) two zoologists.’	X
		Inan	На полке ___ (быть) два мобильника ‘On the shelf ___ (be) two mobile phones.’	X	Красным цветом ___ (писать) два самописца ‘In red ___ (write) (write) five automatic styluses.’
	‘five’	Anim	В классе ___ (быть) пять журналистов ‘In the classroom ___ (be) five journalists.’	X	К своим ___ (пробиваться) пять партизан ‘To their own ___ (get through) five partisans.’
		Inan	X	В коробке ___ (лежать) пять рисунков ‘In the box ___ (lie) five drawings.’	X

Table 3.3: Sample set of stimuli illustrating how the conditions apply. Only quantifiers ‘two’ and ‘five’ are shown.

presentation did not bias participants to give any particular form as a response. Two sample test sentences are given in (17)-(18), while a sample filler sentence is given in (19).

(17) Dva lista __ (rasti) na dereve
Two leaf.GEN __ (grow.INFINITIVE) on tree.LOC
'Two leaves __ (grow) on the tree
(*Quantifier: "two"; Verb: Intransitive; Animacy: Inanimate; Order: SV*)

(18) V bol'nice __ (ostavat's'a) neskol'ko xirurgov
In hospital.LOC __ (remain.INFINITIVE) several surgeon.GEN.PL
'In the hospital there __ (remain) several surgeons.'
(*Quantifier: "several"; Verb: Inversion; Animacy: Animate; Order: VS*)

(19) Otec prines kastr'ul'u so __ (svežij) ikroj
Father.NOM bring.PAST.MASC pan.ACC with __ (fresh.MASC.SG.NOM) caviar.INST.SG
'Father brought a pan with __ (fresh) caviar in it.'
*Intended completion: svežej 'fresh.*FEM.SG.INST*'*

Of the 96 filler sentences, 76 contained gaps corresponding to a noun or adjective, and 20 included gaps for verbs. The fillers with verb-gaps were included so that participants did not learn to associate sentences containing verb-gaps with sentences containing quantified subject NPs. In every filler sentence, the intended form of the given word was unambiguous. All nouns and adjectives were given in citation form, which is the nominative singular, and additionally for adjectives, masculine gender. A list of the fillers is given in Appendix B.

The stimuli for the word-association task consisted of 36 nouns, 36 verbs, 36 adjectives, and 36 adverbs, all given in citation form. These filler words are also given in Appendix B.

3.2.3 Participants

The participants were students at the State University Higher School of Economics in Moscow. All were native Russian speakers, using Russian daily. Fifty-eight participants completed the experiment in exchange for payment that varied depending on the number of questions they answered. Participants who responded to all 288 stimuli (both critical and fillers) received 300 rubles (approximately 10 USD). Participants who elected to quit early received proportionately lower compensation.

3.2.4 Procedure

All tasks were completed over the Internet using SurveyMonkey online survey software. The instructions informed the participants that the study was investigating word choice in different contexts by means of a word-association task. According to the instructions, the sentence-completion tasks were included only as a way of preventing the word-association task of becoming too repetitive. In the sentence completion task, the participants read the sentence on the screen and typed in the form of the word that sounded best, before advancing to the next page to perform the word association task. In this task they were instructed to read the word on the screen and type in the first word that came into their heads.

Although the entire procedure took approximately an hour when it was completed without pause, there was no time limit on any of the tasks. Since the study was performed remotely, it was not uncommon for participants to finish hours or even days after they had started. Upon completion of the study the participants were debriefed and informed of the true purpose — namely, the investigation of influences on singular or plural agreement in sentences with quantified noun phrases. Their consent was then collected one last time. Participants who decided not to finish the study were asked to click an exit button, but more frequently they simply closed the Internet browser window. If they did click the exit button, they were taken directly to the debriefing page, where they could confirm their consent, and their partial data was included in the analysis. If they simply closed their browser window, there was no way to collect their informed consent, and their responses were discarded.

3.2.5 Statistical analysis

Responses were coded for tense (past and non-past), as well as for number (singular and plural). In some cases, participants rewrote the entire sentence to give a different word order, or used a different verb from the one provided. These responses were discarded. The remaining responses were then analyzed using mixed effects logistic regression modeling, with verb number — singular or plural — as the outcome variable and random intercepts for Participant and Word. The potential fixed effects were Quantifier, Verb, Animacy, Order, and Tense, and their interactions.

For factors with more than two levels, previous literature provided estimates of the relative order in which the different levels were expected to condition plural agreement (see Table 3.1). Accordingly, these factors were coded using backwards difference coding. Under this coding system, the levels of a factor are ordered, and the coefficients of the model represent the difference between each level and its immediately preceding level. A positive coefficient indicates a higher probability of observing the outcome variable compared to the immediately preceding level, and a negative coefficient indicates a lower probability.

The analysis was conducted using the R programming environment (R Development Core Team, 2013) with the package languageR (Baayen, 2008). To determine which simple effects should be included, two methods were used: forward entry of predictors, and backward elimination. The forward entry model was built by evaluating the improvement in fit as each individual predictor was included. Predictors were added in the order of the amount of attention they received in the literature, as follows. The starting point was an initial model including only the intercept and random effects for Participant and Word. Its fit was compared to the fit for a model that contained one fixed effect as well as the random effects. Since Animacy is identified most often as a factor that affects agreement realization, it was the first effect added to the baseline model. Order is discussed almost equally as often, while the verb semantics are not mentioned in some sources (e.g. Corbett, 2006), and the identity of the quantifier neglected in others (e.g. Robblee, 1993a). Tense is mentioned only in Kuvšinskaja (2012), and so was added to the model last. The order of addition of predictors was therefore Animacy first, then Order, Verb, Quantifier, and Tense. The addition of each one significantly improved the fit of the model.

To validate the inclusion of each of the five simple effects, the model was then inspected using backward elimination. This method involved taking the full model and testing it against a simpler model created by removing one of the predictors. This validation is necessary because later predictors in a model might equally well explain variability in the data that was explained by a predictor added earlier in the fitting process. For example, Animacy might have significantly improved the fit of the model only because it was the first predictor added to the null model. When later predictors, such as Quantifier or Tense, were added, Animacy might not have contributed anything to the model fit. For this reason, the full simple effects model was simplified by taking out predictors in the same order in which they were added. Each simplified model was tested against the full simple effects model, to determine whether the contribution to model fit associated with the earliest predictors was still present when the later predictors were in the model. As it turned out, even in the presence of all later predictors, each factor still significantly improved the fit of the model, and so the results of the backwards elimination model-building process matched the results of the forward entry method.

After the simple predictors were determined, interactions were added. The empty cells in the design and and the relatively small amount of data for the number of predictors precluded testing all possible interactions. Therefore, the interactions were added through forward entry and validated through backward elimination, as before. Each interaction term was selected initially based on a visual inspection of the data. Figure 3.1 shows two examples of the plots used for the visual inspection. The plot on the left illustrates the comparison of plurality rates for SV and VS Order across all values of Quantifier. According to Robblee (1997)'s data, the effect of word order is much smaller for the numeral 'two' than for the the indefinite quantifiers 'several' and 'few' and numerals greater than four. Contrary to this finding, however, the data in the present study showed no obvious tendency in this direction, and for that reason an interaction term for Order × Quantifier was not included in the model-building process. On the other hand, the plot on the right shows a striking apparent interaction between quantifier and animacy that has not been discussed in previous investigations. Sentences with the quantifier 'several' and, to a lesser extent, 'five,' have a greater proportion of plural agreement for inanimate subjects, contrary to the overall pattern of animate subjects favoring plural agreement. The interaction term for Animacy × Quantifier, therefore, was included in the model. After the addition of each potential interaction term, the new model was tested against the old one to determine whether the interaction improved its fit. When the data would permit no further interactions terms, the existing terms were then removed in the same order in which they were added, to determine whether they still improved the fit of the model in the presence of all the other interactions. As with the simple effects, no interactions that improved model fit during the forward entry process turned out to be unnecessary during backward elimination.

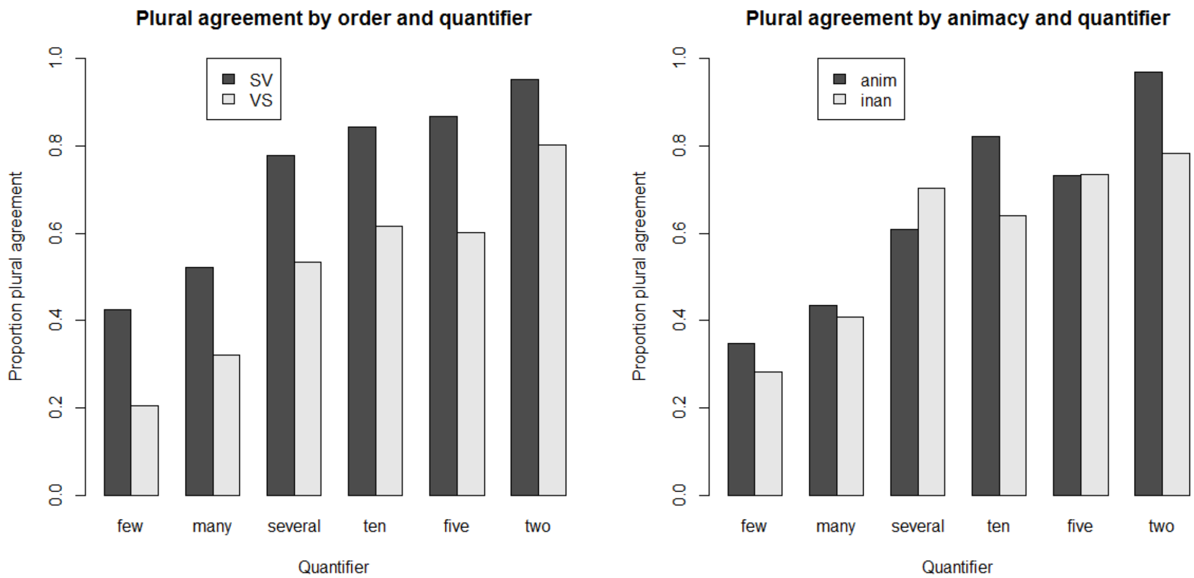


Figure 3.1: Comparisons of the rates of plural agreement across all quantifiers for different word orders (left) and animacies (right).

3.3 Results

Of the 58 participants who were debriefed and gave their final consent, one quit without answering any questions, while a second quit after answering only six critical test stimuli. Of the total 2029 responses, another 29 were discarded because the participants rewrote the whole sentence or gave an entirely different verb from the one provided. The remaining 2000 responses from 57 participants were retained for analysis.

3.3.1 Summary of qualitative patterns

Exactly 59.5% of responses were past tense, while the remaining 40.5% were non-past. In line with Kuvšinskaja (2012)'s findings, subjects did indeed prefer to use past tense for Intransitive and Inversion verbs, as can be seen in Figure 3.2.

The overall proportions of singular and plural agreement were qualitatively similar to previous findings. The left side of Table 3.4 shows the counts of singular and plural responses in the current study according to each of the factors. The right side gives the proportions observed in previous corpus studies. On both sides, the relative rates of plural agreement within a category pattern similarly. For example, the current study found a rate of 49.5% plural agreement for Inversion predicates, which is much higher than the rate of 8.1% observed in Robblee (1993a). Yet in both cases the rate of plural agreement is higher for Intransitive predicates (60.6% here, and 49.7% in Robblee (1993a)) and higher yet for Agentives (76.4% here, and 86.7% in Robblee (1993a)). This

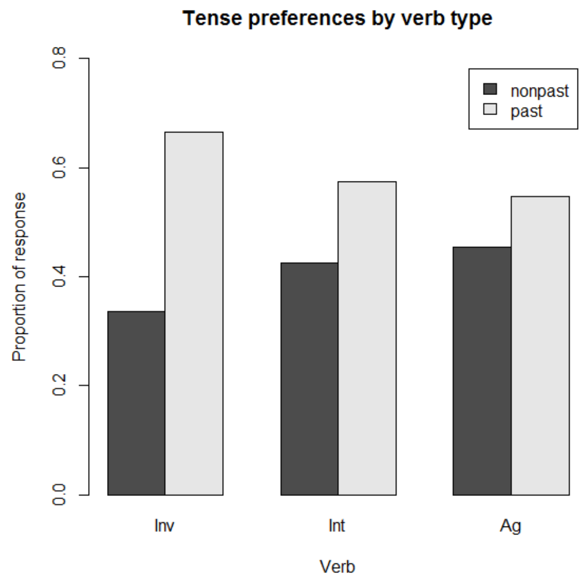


Figure 3.2: Tense choices by verb type. Participants' preference for past tense decreases as verb agentivity increases.

pattern can be seen in all categories: Moving from the topmost level within a category downward, rates of plural agreement tend to increase.

3.3.2 Simple effects

As can be seen in model summary in Table 3.5, all coefficients for simple effects were significantly different from 0 in the predicted direction. Inanimate subjects, Past tense, and VS word order all significantly lowered log odds of plural agreement compared to Animate subjects, Nonpast tense, and SV word order (all $p < 0.001$).

The coefficients of the predictor Quantifier, coded with a backwards difference coding scheme, also confirm the expected effect of Quantifier. Recall that previous studies found that the Quantifiers used here should be ordered as follows, from lowest to highest rates of plural agreement: 'few' < 'many' < 'several' < 'ten' < 'five' < 'two.' The coefficients in the model confirm that the different quantifiers correspond to varying degrees of plural agreement in the expected directions: 'many' more than 'few,' 'several' more than 'many', 'ten' more than 'several', and 'two' more than 'five.'

The effect of Verb was not as straightforward. This study was designed to test Robblee (1993a)'s three categories of verbs, which were previously observed to condition plural agreement in the following order: Inversion < Intransitive < Agentive. The inclusion of the predictor Verb did significantly improve the fit of the model, suggesting that separating verbs into different semantic classes does help predict whether a speaker will choose singular or plural agreement. The

Predictor	Current results Sg	Current results Pl	Current results Sum	Sample results in previous literature Sg	Sample results in previous literature Pl	Sample results in previous literature Sum	Source
Quantifier							
'few'	228 (68.5%)	105 (31.5%)	333	103 (71.0%)	42 (29.0%)	145	Robblee (1997)
'many'	194 (57.9%)	141 (42.1%)	335	212 (96.8%)	7 (3.2%)	219	
'several'	115 (34.4%)	219 (65.6%)	334	138 (64.2%)	77 (35.8%)	215	
'ten'	89 (26.9%)	242 (73.1%)	331	110 (50.0%)	110 (50.0%)	220	
'five'	89 (26.7%)	244 (73.3%)	333				
'two'	41 (12.3%)	293 (87.7%)	334	74 (13.7%)	467 (86.3%)	541	Corbett (1983)
Total	756 (37.8%)	1244 (62.2%)	2000	637 (47.5%)	703 (52.5%)	1340	
Verb							
Inversion	334 (50.5%)	328 (49.5%)	662	113 (91.9%)	10 (8.1%)	123	
Intransitive	265 (39.4%)	407 (60.6%)	672	82 (51.3%)	78 (49.7%)	160	
Agentive	157 (23.6%)	509 (76.4%)	666	12 (13.3%)	78 (86.7%)	90	
Total	756 (37.8%)	1244 (62.2%)	2000	207 (55.5%)	166 (44.5%)	373	Robblee (1993a)
Animacy							
Inanimate	406 (40.9%)	587 (59.1%)	993	1047 (58.6%)	740 (41.4%)	1787	
Animate	350 (34.8%)	657 (65.2%)	1007	790 (37.9%)	1293 (62.1%)	2083	
Total	756 (37.8%)	1244 (62.2%)	2000	1837 (47.5%)	2033 (52.5%)	3870	Patton (1969)
Order							
VS	487 (48.7%)	512 (51.3%)	999	96 (47.3%)	107 (52.7%)	203	
SV	269 (26.9%)	732 (73.6 %)	1001	10 (5.5%)	172 (94.5%)	182	
Total	756 (37.8%)	1244 (62.2%)	2000	106 (27.5%)	179 (72.5%)	385	Corbett (1983)
Tense							
Past	567 (47.6%)	623 (52.4%)	1190	160 (37.0%)	273 (63.0%)	433	
Non-past	189 (23.3%)	621 (76.7%)	810	102 (22.3%)	355 (77.7%)	457	
Total	756 (37.8%)	1244 (62.2%)	2000	262 (29.4%)	628 (70.6%)	890	Kuvšinskaja (2012)

Table 3.4: Observed proportions of singular and plural responses according to Quantifier, Verb, Animacy, Order, and Tense, compared to proportions observed in previous studies. Data from previous studies on the effect of quantifier were a combination of two papers. Robblee (1997) reports the proportions for the quantifier *malo* and *nemalo* together. Corbett (1983) reports the proportions for the other quantifiers, but collapses the responses for numerals between 5 and 10.

fact that there was no significant difference between the adjacent levels of Verb, however, suggests that the current division is not warranted. Therefore, the data were re-coded in order to compare Agentive verbs directly with Inversion verbs. This coding revealed that Agentive verbs *do* differ from Inversion predicates (increase in log odds of 1.89, $p = 0.001$). The effect of Verb is therefore best described by saying that Inversion verbs are associated with significantly lower log odds of plural agreement than Agentive.

3.3.3 Interactions

The model building process described in Section 3.2.5 yields three interaction terms: Quantifier × Animacy, Verb × Tense, and Quantifier × Verb. The improvement in model fit associated with each term confirms the observed non-uniformity of plural agreement rates across Animacy and Quantifier, across Tense and Verb, and across Quantifier and Verb, illustrated graphically in Figure 3.3. In the top left pane, note that the interaction is illustrated only for data with VS order. This is because the study design was not fully crossed, so different combinations of each value of Animacy and Quantifier have different numbers of SV and VS sentences. Order itself influences agreement patterns, so if all the data were illustrated, different bars would represent different numbers of SV and VS Orders. This would obscure the interaction revealed by the model, so Figure 3.3 avoids this problem by graphing only half the data.[3] This top left pane, combined with the model summary, indicates that Inanimate nouns have lower rates of plural agreement when they are quantified by ‘many’ and ‘ten’ than would be expected based on the simple effects of Animacy and Quantifier, and much lower rates when they are quantified by ‘two.’ The top right pane shows that Past tense verbs have lower rates of plural agreement when they are used with Intransitive verbs, and higher rates when they are used with Agentive verbs, compared to the expected rates based on the simple effects of Tense and Verb. And, finally, the bottom panel shows that the increase in plural agreement associated with Intransitive verbs compared to Inversion verbs is enhanced for the Quantifier ‘many,’ but reduced for the Quantifier ‘several.’ In other words, Intransitive verbs increase in plural agreement rates more than Inversion verbs when moving from the quantifier ‘few’ to ‘many,’ but the increase is reduced compared to the increase in Inversion verbs when moving from ‘many’ to ‘several.’

3.3.4 Model fit

Two measures were used to evaluate the fit of the model. The first makes the simplifying assumption that an observation with a fitted value greater than a given threshold is predicted to be plural, while observations with lower fitted vowels are predicted to be singular. The threshold used here was 0.62, which is the overall rate of plural agreement in the responses. In this way, it is possible to simply count up the number of “right” and “wrong” predictions made by the model. An accuracy matrix is given in Table 3.6, showing an overall correct prediction rate of 85.4%.

[3] Indeed, since the pattern for VS and SV word order differ for the interaction of Quantifier and Verb, a logical next step would be to test for a three-way interaction of Order × Quantifier × Verb. Unfortunately, the scarcity of the data and missing conditions in the study design made this impossible.

	Coef β	SE(β)	z	p
Intercept	3.71	0.36	10.3	**<.0001**
Animacy				
Inanimate	−0.96	0.21	−4.5	**<.0001**
Order				
VS	−1.96	0.16	−12.0	**<.0001**
Tense				
Past	−1.73	0.18	−9.5	**<.0001**
Verb				
Inversion vs. Intransitive	0.99	0.56	1.8	0.1
Intransitive vs. Agentive	0.89	0.58	1.5	0.1
Quantifier				
‘few’ vs. ‘many’	1.35	0.33	4.1	**<.0001**
‘many vs. several’	1.62	0.30	5.5	**<.0001**
‘several’ vs. ‘ten’	1.37	0.35	3.9	**<.001**
‘ten’ vs. ‘five’	−0.08	0.09	−1.0	0.3
‘five’ vs. ‘two’	2.74	0.61	4.5	**<.0001**
Quantifier × Animacy				
‘few’ vs. ‘many’ × Inanimate	−1.02	0.47	−2.1	**<.05**
‘many’ vs. ‘several’ × Inanimate	0.37	0.44	0.8	0.4
‘several’ vs. ‘ten’ × Inanimate	−1.09	0.49	−2.2	**<.05**
‘ten’ vs.‘five’ × Inanimate	0.07	0.12	0.6	0.6
‘five’ vs. ‘two’ × Inanimate	−1.56	0.72	−2.2	**<.05**
Verb × Tense				
Inversion vs. Intransitive × Past	−0.81	0.41	−2.0	**<.05**
Intransitive vs. Agentive × Past	1.41	0.41	3.4	**<.001**
Quantifier × Verb				
‘few’ vs. ‘many’ × Inversion vs. Intransitive	1.45	0.58	2.5	**<.05**
‘many’ vs. ‘several’ × Inversion vs. Intransitive	−1.09	0.53	−2.0	**<.05**
‘several’ vs. ‘ten’ × Inversion vs. Intransitive	0.94	0.53	1.8	0.1
‘ten’ vs.‘five’ × Inversion vs. Intransitive	0.05	0.12	0.4	0.7
‘five’ vs. ‘two’ × Inversion vs. Intransitive	−0.94	0.70	−1.4	0.2
‘few’ vs. ‘many’ × Intransitive vs. Agentive	−0.22	0.54	−0.4	0.7
‘many’ vs. ‘several’ × Intransitive vs. Agentive	0.43	0.54	0.8	0.4
‘several’ vs. ‘ten’ × Intransitive vs. Agentive	−0.13	0.63	−0.2	0.8
‘ten’ vs.‘five’ × Intransitive vs. Agentive	−0.19	0.15	−1.3	0.2
‘five’ vs. ‘two’ × Intransitive vs. Agentive	−0.16	0.78	−0.2	0.8

Table 3.5: Summary of the mixed effects logistic regression model. Coefficients for the ordered factors Quantifier and Verb show the difference between two adjacent levels. Positive coefficients indicate that the second level has a higher probability of plural agreement than the preceding one.

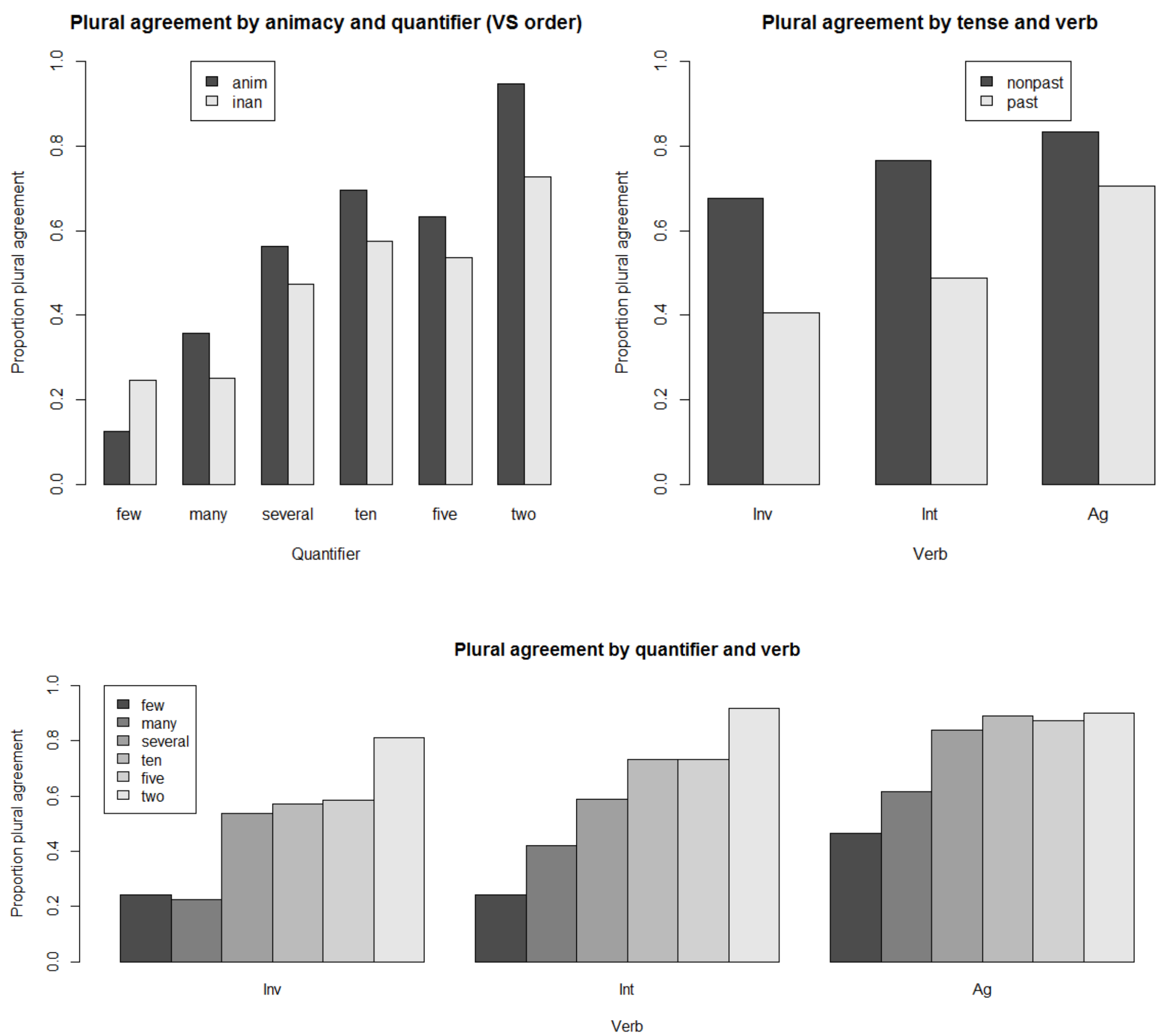


Figure 3.3: Interactions in plural agreement rates between quantifier and animacy (top left; note that only data with VS order is given), tense and verb (top right), and quantifier and verb (bottom).

The second measure of model fit evaluates the specific probabilities generated by the model. Figure 3.4 shows one way of doing this. The predicted probabilities of plural agreement are binned into deciles, and plotted against the observed proportions of plural agreement for each bin (Baayen, 2008). As Figure 3.4 shows, the correlation of predicted and actual proportions of plural agreement is very high ($R^2 = 0.997$).

	Plurality		Total
Accuracy	Singular	Plural	
Right	658	1050	1708
Wrong	98	194	292
Percent Correct	87.0%	91.2%	85.4%

Table 3.6: Accuracy matrix for the mixed effects logistic regression model.

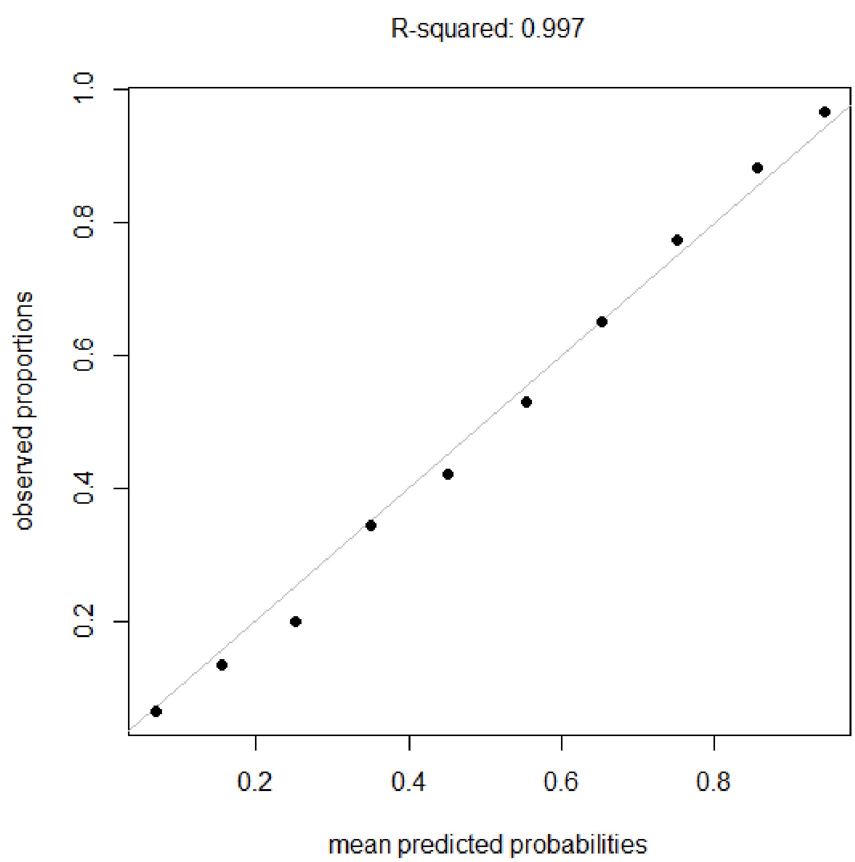


Figure 3.4: A plot of the predicted probabilities against the observed proportion of plural agreement for each decile of the data. A perfect fit would generates points that fall exactly on the gray line.

3.4 Discussion

The experiment presented here was designed to complement and elaborate on the findings of previous corpus studies investigating influences on Russian agreement variation, and the results largely confirmed these previous findings. Even when multiple predictors and interaction terms are included in the same analysis, Animacy, Quantifier, Verb, Tense, and Order still have independent contributions to make to the choice of verb form, largely in the predicted directions: Inanimate subjects, VS word order, and Past tense all lower the probability of observing plural agreement, while the probability of seeing plural agreement increases in the predicted direction along the scale of Quantifiers: 'few' is the lowest, followed by 'many,' 'several,' 'ten' and 'five,' and 'two' is the highest. Indeed, even the absence of a difference between the Quantifiers 'five' and 'ten' is consistent with the general division of numerals into paucal ('two', 'three,' and 'four') and non-paucal (e.g. 'five' through 'ten'). While paucals are expected to differ from non-paucals, the non-paucals themselves are often treated as patterning similarly (e.g. Corbett, 1983), which is exactly the finding here for the Quantifiers 'five' and 'ten.' Similarly, the effect of Verb also revealed a significant increase in the probability of plural agreement for Agentive verbs compared to Inversion verbs. The lack of reliable difference in the pairwise comparisons between Inversion and Intransitive verbs, and between Intransitive and Agentive verbs, suggests that Robblee (1993a)'s three-way distinction is not warranted, but a two-way distinction between activity predicates and existence predicates, such as the one identified by Kuvšinskaja (2012) and Corbett (1983), or between dynamic and non-dynamic predicates, as used in Patton (1969), would certainly be consistent with the distinction between Inversion and Agentive verbs observed here.

These findings add to the existing body of knowledge regarding probabilistic variation in language production (e.g., Bod et al., 2003; Bresnan et al., 2007; Gahl and Garnsey, 2004; Jaeger, 2010; Jurafsky, 2002). One tempting interpretation of these usage patterns would be to classify them with other cases of grammatical variation. As outlined in the Introduction, many grammatical constructions permit variant realizations depending on multiple other properties of the sentence, and people seem to have no difficulty juggling these properties and tracking what is most likely to appear, both in their own speech (e.g., Kuperman and Bresnan, 2012) and in the speech of others (e.g., Bresnan and Ford, 2010). Why should this case of grammatical variation be any different?

An alternative view is proposed in Timberlake (2004): All of the effects of these different factors can be unified under one semantic domain — namely, the amount of individuation of the activity described in the sentence. When the quantified entity is more easily conceptualized as a single unit, the verb is more likely to appear with singular agreement, and when the quantified entity is more easily conceptualized as a set of multiple individuals, the verb is more likely to appear with plural agreement. Consider, for example, the seemingly counterintuitive fact that, as numerical quantifiers express increasingly larger quantities, the verb is increasingly likely to take singular agreement. Corbett (1998) attributes this tendency to the ability of speakers to easily individuate small groups of two or three and understand the actions as the behavior of a set of separate individuals. Larger groups of fifty or a hundred, however, cannot be so easily thought of as individuals. This distinction can be compared to the distinction between a few disgruntled vandals and a mob of riotors: Although both collections are made up of the same individuals, the latter is best conceptualized as a single unit. In Russian, this distinction is responsible for the increased preference for plural agreement associated with larger numerals.

This same reason accounts for the fact that animate nouns — which are more easily understood as agentive entities — also prefer plural agreement. The effect of VS order follows suit: Presentational semantics (usually encoded by VS order) introduces a new entity into the discourse, and it is harder to individuate a new entity than a familiar one. Finally, the effect of tense can also be included in this explanation. Events that took place in the past are more likely to be completed than events which are ongoing, and completed events can be more easily understood as a single activity, rather than an ongoing activity performed by multiple individuals.[4]

Under this individuation account, the factors that have been identified as significant predictors do not themselves influence the choice of agreement relations. The true influence is the degree to which the action can be thought of as distributed among multiple distinct entities. Utterances with a more individuated conceptual message are more likely to be realized with plural agreement. Researchers cannot directly measure the degree of notional plurality associated with the intended message, so they must infer it by examining other properties of the utterance, such as animacy of the subject and the order of the sentence constituents. Yet these properties are not themselves the true predictors that speakers use when selecting a grammatical realization. Rather they consult directly what the researcher cannot see: The amount of individuation in the message component of his utterance.

The current findings have implications for the claims about the role of notional agreement in Russian, compared to other languages. Lorimor et al. (2008) found that, compared to other languages such as English, Dutch, French, Italian, Spanish, and Slovak, Russian consistently appears at the low end of the scale of rates of number agreement attraction, gender agreement attraction, and notional number agreement. They propose that this relative robustness of the agreement system (at least in attraction constructions) is due to the rich morphology of Russian verbal inflection: "The richer the number morphology of a language, the more reliable it is at maintaining grammatical agreement relationships and the less likely it is to succumb to notional agreement" (pg. 790-791). This argument, however, is problematic for two reasons. First, if agreement variation of the type observed here is only influenced by the structural properties of the sentence, then why do those properties include effects such as the degree of individuation of the verb? It cannot simply be explained away as a quirk associated with a particular verb lemma. Even when individual verb lemmas were assigned a random effect in the model described here, the difference between the rates of plural agreement for Agentive and Inversion verbs was significant. If, as it seems to be, the actual *meaning* of the verb influences its agreement patterns, then notional semantics must be able to influence the agreement morphology. The second problem is the consistency of the direction of influence associated with the factors observed here. In every case, the condition that promotes easier individuation at a conceptual level favors plural agreement.[5] Therefore, although other research has shown that speakers do have a very fine-grained knowledge of usage patterns in their language, and although the agreement variation observed here can be quite accurately modeled according to the usage patterns of other parts of the sentence, it is more probably determined by the notional plurality of the message.

[4]If this account is accurate, then it makes the prediction that the true distinction is not actually tense, but aspect: Imperfective aspect should prefer plural agreement, while perfective aspect should prefer singular agreement. Because this study did not control for aspect in the stimuli, it was not possible to test this hypothesis with the data presented here.

[5]This is also true of many of the other properties not investigated here.

3.5 Conclusion

This chapter consisted of the collection and analysis of experimental data on Russian verb agreement with quantified subjects. Multivariate analysis of the five properties previously hypothesized to influence the rates of plural verb agreement with quantified subjects showed that these properties do indeed have independent effects. It is not possible to attribute the effect of animacy, for example, solely to the fact that animate subjects most frequently occur with agentive verbs. Further, the regression model showed three two-way interactions between those properties, providing a subtler understanding of how agreement morphology is associated with other properties of a given sentence. The experimental data here provide complementary confirmation of the corpus data collected in previous studies. Although absolute proportions may not match — hardly surprising, if for no other reason than the vastly different sources of data — the qualitative patterns of the experimental data conform robustly to the qualitative patterns of the corpus data. Whether the realization of the agreement morphology is directly influenced by other properties of the utterance, or simply correlates with them because both reflect the notional individuation of the message, it nevertheless seems clear that the systematic variation can be observed across speakers, in both corpus and experimenal data.

Beyond its theoretical contributions, the current chapter has a more utilitarian value. The model presented in Table 3.5 provides a tool by which the precise probability of observing plural agreement can be calculated. It is therefore by means of this tool that contextual probability is calculated in the following Chapter, in which the pronunciation of plural and singular agreement suffixes will be examined.

Chapter 4

Contrast-dependent pronunciation variation: Evidence from Russian

4.1 Introduction

As we saw in Chapter 2, pronunciation is sensitive to probability, but the nature of this relationship depends on the nature of the probability. When a linguistic unit is more probable in the context of an utterance, it tends to be phonetically reduced. This effect has been observed in many different languages, including English (Chapter 2, Aylett and Turk, 2006; Bell et al., 2009; Gregory et al., 1999; Jurafsky et al., 2001; Lieberman, 1963; van Son and Pols, 2003), Dutch (Schuppler et al., 2012; Tily and Kuperman, 2012), Finnish, and Russian (van Son et al., 2004). By contrast, when a linguistic unit is more probable with respect to other members of its morphological paradigm, it is phonetically enhanced (Chapter 2, Kuperman et al., 2007; Schuppler et al., 2012). Yet what, exactly, counts as phonetic enhancement or reduction?

Existing studies have treated reduction as a type of pronunciation that provides less phonetic information regarding the identity of the linguistic unit. The classic feature that indexes reduction is duration: Shorter units have a reduced pronunciation relative to longer ones. Yet duration is not the sole feature. Indeed, often reduction in duration will appear along with reduction along other features. Gregory et al. (1999), for example, found that coronal stops in English were shorter, more likely to be flapped, and more likely to be omitted entirely when they were in contextually probable words. Van Son and colleagues have found consistently similar reductions in duration and spectral center of gravity (van Son and Pols, 2003; van Son and van Santen, 2005; van Son et al., 2004), and Jurafsky et al. (2001) and Aylett and Turk (2006) found similar evidence that contextually probable linguistic units will have both shorter duration and more centralized vowels.

Yet this coupling between various phonetic features that has been observed in other work was absent in the findings from Chapter 2. In that experiment, the duration and center of gravity of the third-person singular *-s* suffix were analyzed as a function of both contextual probability and paradigmatic probability. Contextually probable suffixes in low-frequency verbs showed shorter duration relative to the duration of the stem, while suffixes with a higher paradigmatic probability showed a longer absolute duration. Yet there was no such relationship between probability and spectral center of gravity. The different behavior of these two features flies in the face of previous work that has found such a close connection between reduction along multiple phonetic features. The goal of this chapter is to probe this dissociation. Is it the case that the experimental design simply had insufficient power to detect an association between center of gravity and suffix probability, or is it the case that only one feature, duration, showed variation? If the latter, then why do only some phonetic features participate in probabilistic variation, and how are those features determined?

The current chapter explores a hypothesis that I shall call Contrast Dependent Pronunciation Variation (CDPV). According to this hypothesis, probabilistic pronunciation variation targets phonetic features which are most salient for distinguishing a particular contrast. Saliency of contrast is determined by which forms compete most strongly for selection during speech production. When multiple forms are competing for selection, they are distinguished from each other by multiple phonetic features, and so phonetic variation will affect all of them. In this situation, no contrast is particularly salient. However, when only a small set of forms are competing, then it is possible to distinguish them on the basis of fewer phonetic features, and so phonetic variation will be confined to those features.

To evaluate CDPV, the current chapter presents an experiment on Russian pronunciation variation in subject-verb agreement suffixes. As in Chapter 2, the experiment will focus on number agreement, and, as in Chapter 2, the pronunciation of these suffixes will be analyzed as a function of probability along two domains. The first is contextual probability: Leaving aside the specific verb hosting the suffix, how likely is the suffix to appear in the given sentence? In other words, how probable is singular agreement or plural agreement in the context of the utterance? The second domain is paradigmatic probability: Leaving aside the context of the utterance, how likely is the specific suffix to appear on the particular verb root? In other words, how probable is the plural or singular form to be used from among the other possible forms in the verb's inflectional paradigm?

Number agreement suffixes in Russian allow the exploration of CDPV for several reasons. First, constructions allowing variable subject-verb number agreement make it possible to constrain the set of forms competing for selection down to exactly two competitors: Singular and plural. Second, both of the suffixes are vocalic, and differ primarily in backness, a feature that was not relevant for the contrast explored in Chapter 2. This means that, in Russian, the phonetic feature that is primarily responsible for distinguishing the plural and neuter singular forms is F2, rather than duration. If it is the case that suffix pronunciation varies on all phonetic features, then the suffixes should vary in duration, F1, and F2. If, however, CDPV is correct, then the primary domain of probabilistic pronunciation variation should be confined to F2 and perhaps F1, because those are the phonetic features that are of primary importance in distinguishing the key morphological forms from each other. A further advantage of Russian is the fact that the singular and plural suffixes have distinct forms, making it possible to examine whether the patterns of pronunciation variation apply equally to both of them.

4.1.1 Key properties of Russian

Much of the work studying probabilistic pronunciation variation has focused on Germanic languages — in particular, English and Dutch. It is therefore worth briefly outlining the important properties of Russian that will come into play in this experiment.

4.1.1.1 Phonology

Russian phonology is characterized by a palatalization contrast that affects most consonants, and a robust system of stress-dependent vowel reduction. In stressed syllables, the language has a system that is alternately described as containing five vowels or six vowels: [i, (ɨ), e, a, o, u] (see Timberlake (2004, section 2.2.2) for a discussion); in unstressed syllables, the system reduces to two or three vowels. The type of reduction depends on the palatalization of the preceding consonant, and the position of the syllble with respect to the stressed syllable. After unpalatalized consonants, the vowel system reduces to [ɪ, ʌ, ʊ] in word-initial or immediately pretonic position, and to [ɪ, ə, ʊ] in post-tonic position. After palatalized consonants, the system reduces to [ɪ, ʊ] in all positions, with the post-tonic realizations closer to schwa than the word-initial or pretonic variants (Timberlake, 2004). In the current study, only stressed vowels are examined, but Padgett and Tabain (2005), in agreement with Barnes (2006), hypothesize that this system of phonological reduction might have arisen from a contraction of the vowel space that arose from the durational differences between stressed and unstressed syllables. If this view is accurate, it is worth briefly reviewing their findings on how the vowel space contracts in unstressed — and therefore shorter — syllables.

Typologically, phonological reduction is overwhelmingly manifested as a neutralization primarily of height contrasts in the vowel space (Barnes, 2006; Flemming, 2005). It is not clear, however, to what extent this pattern is true of Russian reduced vowels. Padgett and Tabain (2005) did observe substantial reduction along the F1 dimension, such that the "floor" of the vowel space is raised, such that the most reduced vowels are no lower than [ə], while the ceiling is also lowered, such that reduced /i/ and /u/ are lower than than their stressed counterparts (Padgett and Tabain, 2005). Yet they also observed effects along the F2 dimension, such that the Euclidean distance between front and back vowels is reduced in unstressed vowels compared to stressed vowels.This difference, however, is less salient than the reduction of F1, and further, is primarily due to fronting of /u/ after palatalized consonants, and backing of /i/ after velarized ones. It is thus possible that reduction in vowel backness is primarily the result of coarticulation arising from the consonantal palatalization contrast, rather than from shrinking the vowel space in shortened vowels. This work therefore suggests that if phonetic reduction will target multiple features in Russian, the primary coupling should be between duration and F1: Shorter vowels should have a reduced F1 range, if palatalization is held constant. By contrast, the CDPV hypothesis holds that phonetic variation should target primarily F2.

There is little research on probabilistic phonetic reduction in Russian, apparently confined to van Son et al. (2004) alone. They observed that stressed vowels which contribute less information to word identification, as well as stressed vowels in more frequent words, show reduction in duration and vowel dispersion, especially in read speech. The nature of this reduction in vowel dispersion, however, cannot be attributed to a particular feature, such as height (F1) or backness (F2), as the analysis defines reduction as the Euclidean distance to a theoretical central position within the vowel space. By analyzing F1 and F2 separately, the current study will therefore narrow down more precisely the domain of phonetic variation.

4.1.1.2 Morphology

Russian has a much richer system of inflectional morphology than English has. Nouns have six cases with both singular and plural forms, and a given adjective agrees with nouns not only in case and number, but also in one of three genders. Even taking into account syncretisms, a noun can have as many as 10 distinct forms, and adjectives up to 13. Russian verbs can have over five times as many: Imperfective verbs in principle can have 67 inflectional forms.[1]

Of particular interest in this work are the forms of the past tense verb. In the singular, the verbs have separate forms for masculine, feminine, and neuter; in the plural, all genders are collapsed. Except for the masculine singular, which has a null suffix, these forms are distinguished through vowel quality, as in (20) below. The current work will focus on the vowels distinguishing the plural form from the neuter singular.

		Singular			Plural
(20)	Infinitive	M	F	N	
	rasti, 'to grow'	ros-∅	rosl-a	rosl-o	rosl-i

4.1.1.3 Syntax

Russian's inflectional system serves a pervasive system of syntactic agreement in all areas of the grammar. Determiners and adjectives agree with nouns in gender, number and case. Relative pronouns and predicate adjectives agree with nouns in gender and number. Verbs agree with nouns in gender and number in the past tense, and in person and number in non-past. For the most part, these agreement relations are invariant. In order to design sentences in which the probability of a particular suffix varies, it is therefore necessary to find a construction that permits multiple agreement realizations.

In Chapter 2, the probability of observing the singular agreement suffix was manipulated through the use of agreement attraction constructions (Eberhard et al., 2005). To parallel the design in Russian, it would therefore make sense to employ a similar construction here. Yet, as Lorimor et al. (2008) observed, Russian speakers are less sensitive to these constructions than are speakers of English, Dutch, Spanish, and Italian. For this reason, a different construction was chosen: Quantified subject noun phrases. As we saw in Chapter 3, sentences with quantified subject noun phrases (e.g., *five boys; several chairs*) allow both plural and neuter singular agreement. The probability of using one or the other form depends on properties such as word order, verb semantics,

[1]This includes the four past and six non-past finite forms, the infinitive, two imperatives, two gerunds, and 52 participial forms (four participles, each with 13 inflectional forms).

quantifier identity, and subject animacy. Accordingly, all critical stimuli in the current experiment contained this type of subject.

4.1.2 Other sources of pronunciation variation

The goal of the current experiment is to probe the nature of pronunciation variation in Russian verbal agreement suffixes. Of particuar interest is that variation which can be attributed to contextual or paradigmatic probability of the given suffix, but that is not the only influence on pronunciation. Prosodic properties of the utterance, the surrounding phonetic context, and ideosyncratic speaker variation have effects on phonetic realization — effects that can be much stronger than probabilistic variables. Accordingly, the analysis presented here accounts for all four types of variation, as follows.

Prosodic properties One of the largest sources of pronunciation variation is speech rate: Faster speech entails shorter segment duration, which has consequences for the articulation of vowels. A baseline association between shorter vowels and more centralized pronunciation in Russian has been observed in both stressed (van Son et al., 2004) and unstressed vowels (Barnes, 2006; Padgett and Tabain, 2005). This source of variation is encoded in the sentential speaking rate preceding the verb, measured as the log-transformation of syllables per second (**lgRate**), local speaking rate of the verb, as measured in segments per second at the verb stem (**lgLocalRate**), and, for analyses of vowel quality, the log-transformed duration of the vowel (**lgVowelDur**).

It has further been observed that prosodic structure of the sentence beyond speaking rate also affects pronunciation. Final segments before prosodic boundaries tend to be longer, whether those boundaries are are the result of disfluent pauses (Shriberg, 1999), or reflect grammatical structure, such as morpheme boundaries (Sugahara and Turk, 2009), word-boundaries (Turk and Shattuck-Hufnagel, 2000), or phrase-boundaries (Beckman and Edwards, 1990). The effect of preceding or following disfluencies was encoded by the factors **disPre** and **disPost**, while the effect of preceding or following pauses was encoded with **PausePre**, and **PausePost**. The morphemes of interest in the current study all follow morpheme boundaries and precede word boundaries, so word-level prosodic boundaries were constant across all observations. However, there was variation at the sentence level. Sentences in which the subject followed the verb contained a phrase boundary after the agreement morpheme, unlike sentences in which the subject preceded the verb. This difference was encoded in the factor **order**.

Phonetic context Coarticulatory processes are a highly important influence on the pronunciation of vowels. It has been known for half a century now that the formant structure of a vowel is affected by the surrounding consonants, with distinct patterns of formant transition associated with different places of articulation (Liberman et al., 1967). These patterns are equally strong in Russian, affecting formant transitions between vowels and adjacent consonants in both F1 and F2 (Purcell, 1979). Further, in Russian in particular, a palatalized consonant will produce a strong increase in F2 transitions of any adjacent vowel, as well as a decrease in F1 (Halle, 1959; Purcell, 1979). In this work, all vowels of interest are preceded by the past-tense marker /l/, which is palatalized to $[l^j]$ before plural [i]. This effectively controls the place of articulation and palatalization of the preceding consonant, so controls for phonetic context were necessary only for the following consonants. These were implement by means of the predictors **poaPost**, which recorded whether the following consonant had alveolar, bilabial, palatal, or velar place of articulation, and **palPost**, which recorded whether or not the following consonant was palatalized. Additionally, the factor **vowel** recorded whether a given observation was an observation of the plural suffix *-i* or the singular suffix *-o*.

Speaker variation People would not be able to recognize the voice of an acquaintance if there were not ideosyncratic properties that distinguish the speech of one person from another. These properties can include such patterns as lower or higher voices, or longer or shorter vocal tracts, which can systematically affect pronunciation. Accordingly, the analysis took into account these speaker-dependent variables by adding a speaker-specific random adjustment, and when the structure of the data permitted, a speaker-specific slope.

Probabilistic variation The current study is focused on pronunciation variation that can be attributed to the probability of the linguistic unit of interest — here, the agreement suffix. First is word-form frequency (**lgVerbFq**), which has been shown to be negatively correlated with duration of content words (Bell et al., 2009; Gahl, 2008), as well as duration and centralization of stressed vowels in Russian (van Son et al., 2004). Next are the three measures of probability of interest here. First is contextual probability of observing the particular agreement suffix, operationalized as the logit-transformed probability of plural agreement in the sentence (**logitPred**). The probability of plural agreement was calculated according to the model described in Chapter 3, and the logit-transformation was necessary to create an unbounded numeric predictor from the probability measure, which was bounded at 0 and 1.

The last two probabilistic predictors describe paradigmatic probability. They are the log-odds of using a given suffix over to its competitor, and the inflectional entropy of the entire verb paradigm. The log odds of using a given suffix over its competitor were calculated here as in Chapter 2. Specifically, they are the log-ratio of the frequency the form that was produced — e.g., singular *-o* — to the frequency of the competitor form —- e.g., plural *-i* (**lgRatioFull**).

Inflectional entropy (**entropy**) was calculated as described by Equation 1.6 in Chapter 1. In Chapter 2, we saw that inflectional entropy had no effect on suffix pronunciation, but it is possible that the more complex inflectional paradigms in Russian might permit an effect of entropy to emerge. This calculation is slightly more complicated in Russian than in English, however, because not all verbs share all forms. This problem is particularly evident in the participial system: Perfective verbs can form past participles and past verbal adverbs, but not the present forms, while imperfective verbs can form both. Some verbs, in principle, could have a given participial form, but in practice are never used in that form. This makes it extremely tricky to determine the full range of possible verbal forms a given verb can take. For this reason, the frequencies for a subset of all possible forms were extracted from the RNC: The four past tense forms, the six non-past

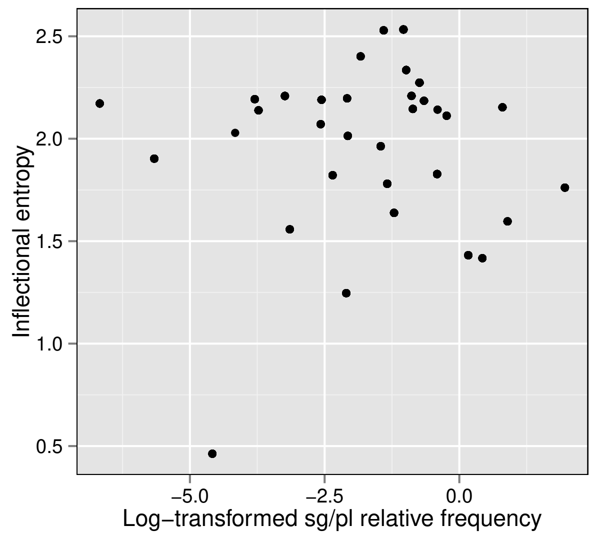


Figure 4.1: Sg-pl entropy measures plotted against relative frequency of the verb forms used in the stimuli.

forms, the infinitive, the two imperatives, and then all participles lumped together, and both adverb forms lumped together. The decision to lump together the participles was further reinforced by the fact that participles behave very similarly to adjectives. They participate in case alternations like adjectives, and all of their inflectional suffixes are purely adjectival. Therefore, it is not even clear that they should be expected to influence questions of processing that involve verbal inflectional paradigms.

A second difference between Russian and English inflectional entropy is the fact that there is practically no correlation between the singular/plural relative frequency and the inflectional entropy of a given verb. In English, the frequency of the singular form is almost always less than the frequency of the plural form, so an increase in relative frequency means that their frequencies become more similar, increasing inflectional entropy. By contrast, in Russian there are so many forms contributing to inflectional entropy that an increase in the relative frequency of just two of them does not have much of an effect on the overall similarity of frequencies across the entire paradigm. This orthogonality between singular/plural relative frequency and inflectional entropy is illustrated in Figure 4.1, which shows the entropy of the singular verbs plotted against their relative frequency. Note that only the pattern for the singular verb forms is shown. The plural verb forms have identical entropies as their singular counterparts, but opposite log-tranformed relative frequencies, so including them in the graph would only create a mirror image. An analysis of the singular variant ratios reveals that there is no correlation between the two measures of paradigmatic probability ($\rho = 0.04$, $p = 0.14$).

A summary of all predictors is given in Table 4.1.

Predictor	Description	Factor levels
PROSODIC		
lgRate	Log-transformed speaking rate in syllables per second	
lgLocalRate	Log-transformed speaking rate of the verb in segments per second	
lgVowelDur	Log-transformed vowel duration	
order	Order of subject with respect to verb	*SV, VS*
PausePre	Pause before verb?	*no, yes*
PausePost	Pause after verb?	*no, yes*
disPre	Presence of perverbal disfluency	*no, yes*
disPost	Presence of postverbal disfluency	*no, yes*
PHONETIC CONTEXT		
vowel	Which vowel was produced	*-i, -o*
poaPost	Following consonant place of articulation	*alveolar, bilabial, palatal, velar*
palPost	Was following consonant palatalized	*no, yes*
PROBABILISTIC		
lgFqFull	Log-transformed verb frequency	
logitPred	Logistic transform of the probability of using *plural* agreement in the sentence	
lgRatio	Log-transformed ratio of used form (e.g., sg) to alternative form (e.g., pl)	
entropyFull	Inflectional entropy of the verbal paradigm	

Table 4.1: Summary of the predictors considered during the model-building process.

4.2 Methods

4.2.1 Participants

Forty-four students and other members of the community surrounding State University Higher School of Economics in Moscow participated in the experiment (34 female and 10 male). All were native speakers of Russian, speaking and using Russian every day.

4.2.2 Materials

Sentences The experimental stimuli consisted of 76 sentences, which were constructed according to the model developed in the norming study. The predicted probability of observing plural agreement in these sentences ranged from 0.025 to 0.998, as illustrated in Figure 4.2. Two counterbalanced lists were created, such that if a sentence contained singular agreement in one list, it contained plural agreement in the other. One list, provided in Appendix C, contained 39 plural

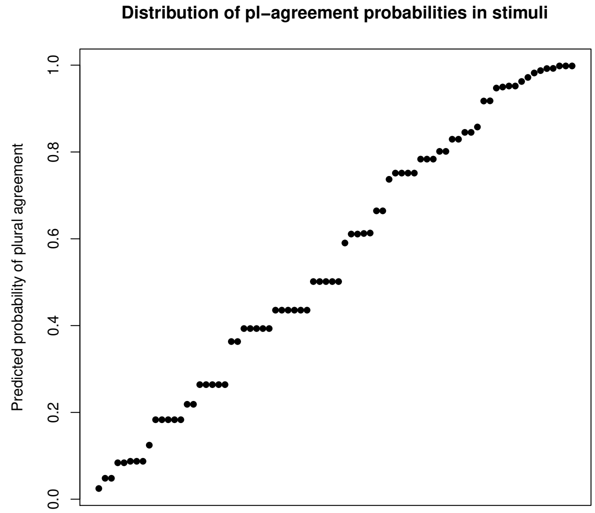


Figure 4.2: Sentence stimuli spanning the full range of predicted probabilities of plural agreement.

verbs and 37 singular, while the other contained 37 plural and 39 singular. As shown in Table 4.2, the number of singular and plural verbs were balanced by probability decade and verb lemma as much as possible.

Every sentence contained four main constituents: A subject, a verb, and two other items, such as an object and a PP-adverbial, or two adverbial phrases. In each sentence the verb was immediately followed by a word starting either in a stop or a nasal, in order to facilitate segmentation. Each sentence contained a gap in either the object or the adverbial phrase, followed by a word in parentheses. The word in parentheses was given in citation form, and participants were instructed to determine the necessary form of the word, and then read the full sentence aloud. This task thus required participants to process the meaning and structure of the sentence, while simultaneously distracting their attention from the verb. Sample stimuli are given below in (21-23).

(21) Na __ (sobranija) vsegda š-l-i mnogo sporov
at __ (meetings) always walk.-PAST-PL many argument.GEN.PL
'Many arguments would always go on at the meeting(s).' P(pl): 0.048

(22) Mnogo tovarov privlek-l-i pokupatelej v __ (magaziny)
many goods.GEN.PL attract-PAST-PL buyers.ACC.PL in __ (stores)
'A lot of merchandise attracted buyers in the stores.' P(pl): 0.264

(23) Dva matrosa volok-l-i dyr'avyj jalik po __ (pesok)
two sailor.GEN.SG drag-PAST-PL hole.ridden dinghy along __
'Two sailors dragged the hole-ridden dinghy along the sand.' P(pl): 0.998

Decade	Number Pl	Sg
0.0–0.09	4	4
0.1–0.19	3	3
0.2-0.29	4	3
0.3–0.39	4	3
0.4–0.49	3	3
0.5–0.59	3	3
0.6–0.69	3	3
0.7–0.79	4	4
0.8–0.89	3	4
0.9–1.00	8	7

(a) Agreement breakdown by probability decade

Lemma	Pl	Sg	Lemma	Pl	Sg	Lemma	Pl	Sg
‘trudge’	1	1	‘sweep’	2	1	‘bring$_a$’	1	1
‘convey’	1	1	‘carry’	1	1	‘bring$_b$’	0	
‘lead’	0	1	‘shephard’	1	0	‘attract’	2	2
‘draw’	0	1	‘transfer’	0	1	‘pronounce’	1	2
‘drag’	1	0	‘bake’	1	0	‘read’	1	0
‘grow up’	0	1	‘weave’	1	1	‘distract’	2	2
‘row’	1	1	‘go’	1	1	‘grow’	4	3
‘burn’	1	0	‘crawl’	1	1	‘scrape’	1	0
‘walk’	3	4	‘help’	0	1	‘rescue’	1	0
‘lie down’	2	1	‘disdain’	1	1	‘guard’	1	1
‘flow’	2	3	‘pound’	0	1	‘shake’	1	0
‘crawl away’	1	0	‘bloom’	2	2			

(b) Agreement breakdown by verb lemma

Table 4.2: As much as possible, verbs occurred equally often in plural and singular forms within each probability decade and for each verb lemma.

In addition to the 76 critical sentences, each list contained the 120 filler sentences, in which the missing word was either a noun (46), verb (43), or adjective (31). These sentences were randomly interspersed with the critical sentences, creating two stimulus lists of 196 sentences each. All fillers are given in Appendix C.

Verbs Thirty-five verbs were used, in two forms each, for a total of 70 distinct verb forms, all end-stressed in the past tense. Because there are not a large number of such verbs in Russian, it was not possible to control length, and some verbs were repeated, up to a maximum of seven times. The frequencies were taken from the Russian National Corpus twice—once from the full corpus, and once from the disambiguated subsection. Using the disambiguated subcorpus had the advantage that it was possible to be sure that every use of a verb form was truly the verbal meaning used in this experiment, rather than an accidentally homographic word with an entirely different usage. For example, the word form *peč’* is both the infinitive of the verb ‘bake,’ but also a feminine singular noun meaning ‘oven.’ Despite this advantage, however, the disambiguated subcorpus, at roughly 5.9 million words, is much smaller than the full corpus of 230 million. As a result there are frequently cases where a particular verb form simply does not occur, and hence it is difficult to get an accurate frequency estimate. Fortunately, the correlation between the logged frequencies of both the verb forms and the verb lemmas in the full and disambiguated sections of the corpus was extremely high, (cor(verb form)=0.996, cor(verbLemma)=0.993, $p < 0.001$), so the full corpus was used for all further lexical measurements.

4.2.3 Procedure

Participants were tested in a private room, using the OpenSesame experimental software (Mathôt et al., 2012) run on a Lenovo T400 notebook computer. Their speech was recorded at a sampling rate of 44,100 Hz, through one of two microphones: A head-mounted microphone attached to an Icicle pre-amplifier, or a Logitech USB Desktop microphone. Sentences were presented for 10 seconds, changing automatically as the experiment continued. Occasionally subjects requested the experimenter to remain in the room with them, and one subject was unable to complete the sentences within 10 seconds, and so restarted the experiment with a 15-second stimulus presentation. Each list was presented in 5 blocks of 40 sentences each, except for the last block, which had 36 sentences. Participants were encouraged to rest between blocks as long as they wanted, and only to resume the experiment when they were ready. The experiment took about 45 minutes in total, and participants were compensated 230 rubles (~$7.50 USD) for participation.

4.2.4 Coding of results

Sentences were divided into five sections: The preamble, or the portion before the verb; the verb stem; the past tense *-l* suffix; the agreement suffix; and the postamble, or the portion after the verb. Acoustic measurements were extracted using Praat (Boersma and Weenink, 2009). All transitions were annotated at the zero crossing in the waveform closest to their respective cues, as follows. Transitions from the *-l* into the vowel were annotated at one of two spectral cues: Either a burst-like transition (most frequent before the plural *-i* suffix) or a sharp increase in amplitude in the waveform (more frequent with the singular *-o* suffix). Transitions from the vowel into a following nasal were annotated at the sharp decrease in amplitude in the waveform, and transitions from the vowel into a following stop were annotated at the offset of F2 in the spectrogram.

The sentence speaking rate (**lgRate**) was calculated as the log ratio of the number of syllables in the preamble to the duration of the preamble, yielding a rate in syllables per second. The local speaking rate (**lgLocalRate**) was calculated as the log ratio of the number of segments in the verb stem to the duration of the verb stem, yielding a rate in segments per second. Pauses within words and errors that were quickly corrected were marked as disfluencies, and coded based on their distance from the verb in units of stressed syllables. For example, in the sentence 24 below, one speaker started to use the preposition *vozle*, 'beside,' instead of the correct *posle*. This occurred in a position five stressed syllables before the verb, so the value for the predictor **disPre** was position -5. A similar procedure was used to code the value for **disPost**. All sentences with a disfluency occurring at the verb—i.e., at position 0—were discarded.

(24)	Pósle	štórma	__	(razbítyj)	stekló	i	músor	**me-l-ó**	mnógo volontérov
	After	storm	__	(smashed)	glass	and	trash	**sweep-PST-N.SG**	many volunteers

'After a storm the smashed glass and trash would be swept up by many volunteers.'

Because there were so few disfluencies either before or after the verb, the predictors **disPre** and **disPost** were collapsed in factors. Rather than indicating the position of a disfluency, **disPre** simply coded the presence of disfluency before the verb, while **disPost** did the same for disfluencies after the verb.

4.2.5 Analysis

Phonetic variation was analyzed with mixed effects regression, using the `lme4` package in the R programming environment (Bates et al., 2013; R Development Core Team, 2013). In each case, the model was built using forward addition. Each predictor was added in two stages: First the model was fit containing the predictor of interest. Next, an interaction was added between the predictor of interest and **vowel**. This was necessary because the inherent properties of one vowel might make it more sensitive to a given predictor than another. A vowel with a relatively low F2, such as *-o*, for example, would be expected to show a much stronger increase in F2 before a palatalized consonant than a vowel like *-i*, whose F2 is already quite high. The improvement in model fit was ascertained by means of a log-likelihood ratio test, and a predictor was retained if it significantly improved the fit of the model either alone *or* when interacting with **vowel**. The final step in the model-building procedure was the addition of the probabilistic predictors of interest, **logitPred**, **lgRatioFull** or **entropyFull**. Their contribution to the control model fit was evaluated after an interaction with **vowel**, parallel to the procedure for the control predictors.

Three phonetic properties were analyzed according to the model-building procedure described above: Duration, F1 and F2. F1 and F2 were calculated as the mean value from the first quarter, midpoint, and third quarter of each vowel token's duration. Due to differences in the anatomy of individual speakers' vocal tracts, and differences in the spectral properties characteristic of *-i* and *-o*, F1 and F2 were *z*-score normalized by speaker and by vowel. This was accomplished by subtracting from each F1 or F2 value the mean F1 or F2 for the given speaker's pronunciation of that vowel, and then dividing by the standard deviation of all F1 or F2 values for the speaker's pronunciation of that vowel.

Random intercepts included verb stem and sentence. In the duration analysis, there was also a random intercept for subject. In the F1 and F2 analyses, however, the normalization procedure took into account individual variation in mean values for each subject, and so the random intercepts did not include a term for subject. Random slopes were also considered, but the complexity of the models did not permit a maximal random effects structure (Barr et al., 2013). For this reason, the contribution of random slopes to model fit was evaluated by means of a log-likelihood ratio test against the full model containing random intercepts only. Only those slopes that created the largest drop in AIC were then included in the model. Usually only one slope could be added before the model failed to converge.

4.3 Results

In total, speakers produced 3344 sentences. Of these 533 were discarded. In most cases (317) the discard was due to a disfluency at the verb, or prosodic evidence that the speaker had misinterpreted the sentence. A further 189 were discarded due to an error somewhere else in the sentence, and a final 28 were discarded due to recording problems. The remaining 2810 observations were retained for analysis.

4.3.1 Duration

4.3.1.1 Model summary

In the duration analysis, the dependent variable was the log-transformed duration of the suffix vowel. A summary of the effects of the control predictors that significantly improved the fit of the model is given in Table 4.3, along with the decrease in AIC and the result of the log-likelihood ratio test. Figure 4.3 illustrates the partial effects of these predictors.

Prosodic predictors mostly showed the expected effects: Vowel duration decreased with faster speaking rate, although the effect was stronger for *-i* than for *-o*. Vowels before pauses were longer, although, again, the effect of a following pause was stronger for *-i* than for *-o*. Finally, vowels were shorter if there had been a disfluency previously in the sentence, although the effect again was mostly restricted to *-i*. This speed-up is perhaps because speakers felt that the early disfluency had delayed their production of the sentence, and spoke faster to make up for lost time after recovering.

Articulatory predictors likewise held few surprises. Singular *-o* was overall longer than plural *-i*, a natural consequence of being a lower vowel (Lehiste, 1976). Further, vowels were shorter before bilabial and velar following consonants. This is probably a consequence of the fact that it is easier to move an articulator that is relatively independent of the tongue tip — such as the lips or tongue dorsum — into position for the following consonant, than it is to move the tongue tip from the vowel-articulating position to the following consonant position. The faster an articulator can be moved into position for a following consonant, the sooner that consonant can be articulated, and the shorter preceding vowels will be. This effect was slightly stronger for *-i* than for *-o*.

The random slopes that best improved the model were the interaction of **vowel** with **PausePost** by **speaker**. The addition of the critical predictors — **logitPred**, **lgRatio**, and **entropyFull** — did not significantly improve model fit, even when they were considered in interaction with **vowel** (lines 14a-15c in Table 4.3). In order to determine whether this lack of fit was due to outlier observations, the data set was trimmed to remove observations whose scaled residuals in the model containing control predictors were greater than 2.5 standard deviations from 0 (Baayen, 2008). This resulted in the loss of 67 observations, or about 2.4% of the data. The model containing only control predictors was refit to this data set (Line Refit 13 in Table 4.4). This is the model whose partial effects are plotted in Figure 4.3. The critical predictors were once again evaluated against this refit control model (Table 4.4). Again, they did not significantly improve model fit.

4.3.1.2 Model evaluation

Figure 4.4 shows two measures evaluating the fit of the trimmed model containing only control predictors and the random slope (line Refit 13 in Table 4.4). The panel on the left shows the predicted vowel duration plotted as a function of the actual vowel duration. The roughly constant

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p
0	**Intercept**	-1.23	0.09	-13.57	1534		
1	**vowel**				1334 (-200)	202.46	< 0.001
	-o	-0.10	0.08	-1.21			
2	**lgRate**	-0.33	0.03	-9.52	1185 (-149)	151.22	< 0.001
3	**vowel:lgRate**				1180 (-5)	6.13	< 0.05
	-o	0.08	0.04	2.00			
4	**lgLocalRate**	-0.45	0.04	-11.37	944 (-236)	238.25	< 0.001
5	**vowel:lgLocalRate**				936 (-8)	9.89	< 0.01
	-o	0.07	0.04	1.78			
6	**PausePost**				84 (-852)	854.31	< 0.001
	yes	0.65	0.02	26.05			
7	**vowel:PausePost**				79 (-5)	6.67	< 0.01
	-o, yes	-0.10	0.04	-2.65			
8	**poaPost**				59 (-20)	26.44	< 0.001
	bilabial	-0.18	0.03	-5.85			
	palatal	-0.10	0.08	-1.29			
	velar	-0.27	0.07	-3.90			
9	**vowel:poaPost**				52 (-7)	12.57	< 0.01
	-o, bilabial	0.05	0.02	2.53			
	-o, palatal	0.001	0.05	0.03			
	-o, velar	0.13	0.05	2.79			
10	**palPost**				47 (-5)	6.98	< .01
	yes	0.11	0.03	3.38			
11	**vowel:palPost**				42 (-5)	6.86	< .01
	-o, yes	-0.05	0.02	-2.64			
12	**disPre**				35 (-9)	9.72	< .01
	yes	-0.16	0.05	-3.63			
13	**vowel:disPre**				33 (-2)	4.04	< .05
	-o, yes	0.12	0.06	2.01			
14a	**lgRatio**	0.003	0.004	0.80	35 (+2)	0.05	0.83
15a	**vowel:lgRatio**				36 (+3)	0.85	0.65
	-o	-0.01	0.01	-0.91			
14b	**entropyFull**	-0.04	0.04	-0.96	34 (+1)	0.51	0.47
15b	**vowel:entropyFull**				35 (+2)	1.34	0.51
	-o	0.02	0.02	0.91			
14c	**logitPred**	0.003	0.01	0.44	34 (+1)	0.60	0.44
15c	**vowel:logitPred**				34 (+1)	3.06	0.22
	-o	0.01	0.005	1.57			

Table 4.3: Summary of duration model. If simple effects do not significantly improve the model, then the effect's interaction with **vowel** is evaluated against the preceding model without either the effect or the interaction. For example, model 14a does not improve fit over model 13, so model 14b is evaluated against model 13, rather than against model 14a.

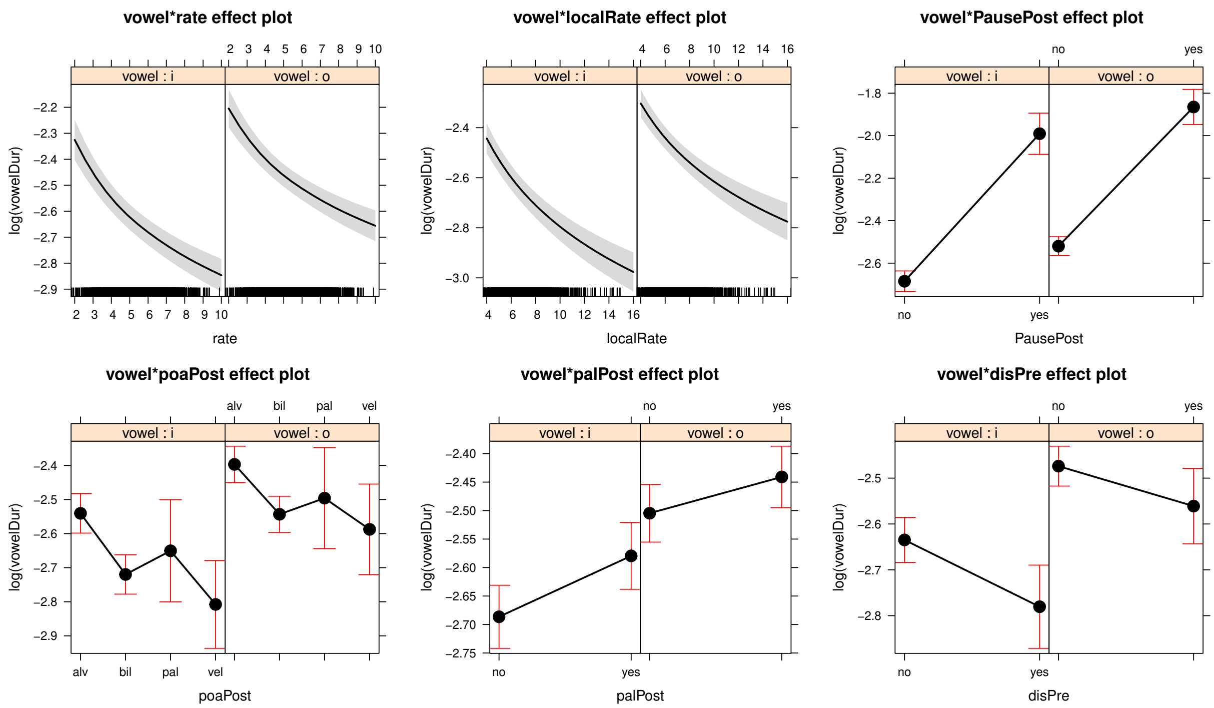


Figure 4.3: Partial effects plot for the control predictors in the duration model.

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p
Refit 13	**vowel:PausePost** by **speaker**				-745		
Refit 14a	**lgRatio**	0.003	0.004	0.70	-743 (+2)	0.20	0.66
Refit 15a	**vowel:lgRatio**				-742 (+3)	0.49	0.78
	-o	-0.003	0.01	-0.54			
Refit 14b	**entropyFull**	-0.03	0.04	-0.98	-744 (+1)	0.46	0.50
Refit 15b	**vowel:entropyFull**				-743 (+2)	1.86	0.39
	-o	0.02	0.02	1.19			
Refit 14c	**logitPred**	0.00	0.01	-0.035	-743 (+2)	0.32	0.57
Refit 15c	**vowel:logitPred**				-747 (-2)	5.57	0.06
	-o	0.01	0.004	2.30			

Table 4.4: Summary of the effects of key predictors when the model includes random slopes and has been trimmed to removed outliers. Improvement in model fit is evaluated as in Table 4.3.

spread of the points shows that the model does not tend to be more accurate for one subset of the data compared to another. This homogeneity is a sign of good model fit. A second sign of good model fit is normality of the residuals. The right panel of Figure 4.4 shows a qq-plot, illustrating that the residuals depart from normality only for the very highest and lowest vowel durations. This is enough for a Shapiro-Wilk test for normality to reach significance, indicating non-normality, but only barely (W=0.999, p =0.046).

The conditional R^2 of the full control model (line Refit 13 in Table 4.4) is 0.73, indicating that almost three quarters of the variance in the suffix duration is explained by the fixed and random effects combined (Nakagawa and Schielzeth, 2013). The marginal R^2 of the control model is 0.5170, indicating that just over half of the variance in suffix duration, or more than two thirds of the explained variance, is accounted for by the fixed effects. The marginal R^2 must be reported to four significant figures in order to highlight the fact that it changes only minimally when the critical predictors are added: Even in interaction with **vowel**, the addition of **lgRatio** brings the marginal R^2 to 0.5174, **entropyFull** brings it to 0.5177, and **logitPred** brings it to 0.5174. In other words, the critical predictors explain at most less than 0.1% of the variance in suffix duration. It is not hard to see why they did not improve model fit.

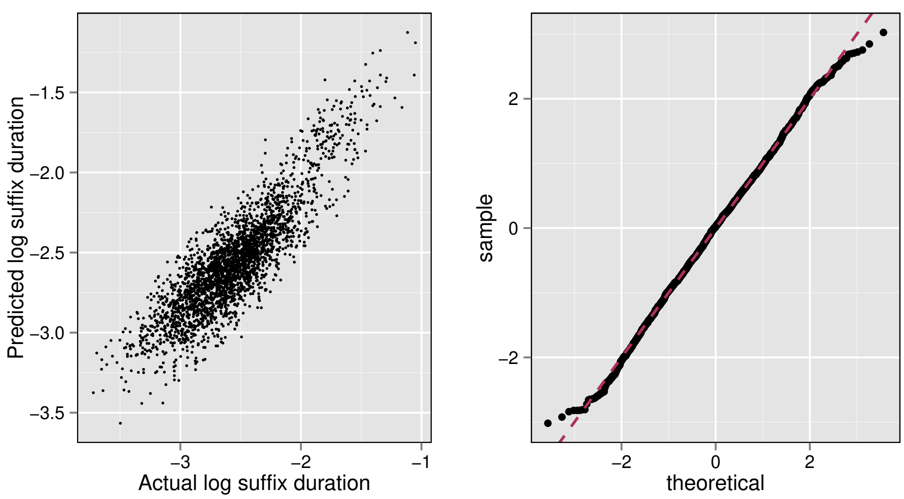


Figure 4.4: Model evaluation plots of the suffix duration model. The left panel illustrates the homogeneity of the residuals, and the right plot shows near-normality.

4.3.2 F1

4.3.2.1 Model summary

Normalized F1 values were analyzed in a manner parallel to the duration analysis, yielding the model summarized in Table 4.5. Prosodic predictors that significantly improved the model fit were vowel duration (**lgVowelDur**), sentence speaking rate (**lgRate**), and local speaking rate (**lgLocalRate**). These had different effects on the different vowels. A longer vowel increased F1 for *-o*, but not for *-i*, and the same was true for a faster sentence speaking rate. By contrast, a faster local speaking rate increased F1 for *-i*, but not for *-o*. These effects are pictured in the top row of Figure 4.5.

Phonetic context likewise interacted with vowel quality. Generally, *-i* was more sensitive to the place of articulation of following consonants (**poaPost**) than *-o*, such that F1 dropped when the final consonants were palatal, and rose with a following velar consonant. The palatalization of the following consonant (**palPost**) also had a stronger effect on *-i* than on *-o*, such that *-i* had a much greater reduction in F1 before a palatalized consonant than *-o* did. These effects are shown in the second row of Figure 4.5.

Even in combination with **vowel**, paradigmatic probability did not affect F1 (**lgRatio**: $\chi^2 = 3.00, p = 0.22$; **entropyFull**: $\chi^2 = 0.10, p = 0.95$). Their effects are shown in lines 12a-13b in Table 4.5. Contextual probability (**logitPred**), however, did significantly improve the fit of the model when it interacted with **vowel** (line 13c in Table 4.5; $\chi^2 = 10.12, p < .01$). The effect was carried largely by the behavior of the singular *-o* suffix. As **logitPred**, or the probability of plural agreement, increased, F1 for *-o* increased ($\beta = 0.06, SE(\beta) = 0.02, t = 3.15$). Put differently, as the probability of singular agreement increased, F1 decreased, resulting in a higher articulation of the singular suffix *-o*.

The only random slope that improved model fit was **logitPred** by **speaker**. Since F1 values were already normalized by speaker, this random slope was added without including any corresponding by-speaker intercept. Next, observations with scaled residuals greater than 2.5 standard deviations from 0 were removed. The control model and the model containing the interaction between **logitPred** and **vowel** were then refit to this trimmed data set. These refinements only strengthened the improvement in fit ($\chi^2 = 16.31, p < 0.001$). Further, in the trimmed model both *-i* and *-o* showed an effect of **logitPred**. With higher **logitPred**, or higher probability of plural agreement, F1 in *-i* decreased, while F1 in *-o* increased. In other words, as the probability of saying a particular suffix increased, the articulation of that suffix became higher, whether it was *-i* or *-o*. It is the effects from this trimmed model that are pictured in Figure 4.5.

4.3.2.2 Model evaluation

Unlike the duration model, the model predicting F1 had a poor fit to the data. The left plot of Figure 4.6 shows the predicted normalized F1 values plotted as a function of the actual values from the model given in the bottom row of Table 4.6. Although the spread of points does not much vary across different values of normalized F1, the spread is consistently wide. The right panel, showing a qq-plot of the residuals, also reveals that the distribution of the residuals does depart from normality at both extreme ends of the data. A Shapiro-Wilk test for normality confirms this departure ($W = 0.9974, p < 0.001$).

The poor fit of the data is reflected in the R^2 of the model. Without any key predictors (line Refit 11 in Table 4.6), the conditional R^2 is only 0.16, indicating that both fixed and random effects together account for a very small portion of the variance in F1, and the marginal R^2 of 0.093 indicates that less than a tenth of that variance is explained by the fixed effects of the model

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p
0	**Intercept**	-0.47	0.26	-1.81	7727		
1	**vowel**				7729 (2)	0.0001	0.99
	-o	0.84	0.34	2.51			
2	**lgVowelDur**	0.07	0.07	0.91	7730 (1)	1.23	0.27
3	**vowel:lgVowelDur**				7715 (-12)	17.88	< 0.001
	-o	0.29	0.11	2.64			
4	**lgRate**	-0.01	0.12	-0.08	7711 (-4)	5.85	< 0.05
5	**vowel:lgRate**				7706 (-5)	6.66	< 0.01
	-o	0.36	0.16	2.29			
6	**lgLocalrate**	0.47	0.14	3.37	7703 (-3)	5.65	< 0.05
7	**vowel:lgLocalRate**				7699 (-4)	5.87	< 0.05
	-o	-0.41	0.17	-2.37			
8	**poaPost**				7692 (-7)	13.31	< 0.01
	bilabial	-0.002	0.08	-0.03			
	palatal	-0.69	0.21	-3.36			
	velar	0.52	0.19	2.72			
9	**vowel:poaPost**				7674 (-18)	23.54	< 0.001
	-o, bilabial	-0.15	0.08	-2.03			
	-o, palatal	0.21	0.19	1.09			
	-o, velar	-0.87	0.19	-4.49			
10	**palPost**				7652 (-22)	23.91	< 0.001
	yes	-0.49	0.08	-6.23			
11	**vowel:palPost**				7642 (-10)	12.74	< 0.001
	-o, yes	0.26	0.07	3.58			
12a	**lgRatio**	-0.02	0.02	-1.45	7641 (-1)	2.91	0.09
13a	**vowel:lgRatio**				7643 (+2)	0.09	0.76
	-o	0.01	0.02	0.31			
12b	**entropyFull**	-0.02	0.09	-0.20	7644 (+2)	0.00	0.95
13b	**vowel:entropyFull**				7646(+2)	0.09	0.76
	-o	0.03	0.09	0.31			
12c	**logitPred**	-0.03	0.02	-1.37	7644 (2)	0.19	0.66
13c	**vowel:logitPred**				7636 (-6)	10.12	< 0.01
	-o	0.06	0.02	3.15			

Table 4.5: Summary of the full F1 model. The predictors **lgVowelDur** and **logitPred** did not significantly improve the model except in interaction with **vowel**. For this reason, the improvement in model fit contributed by these interaction terms is evaluated in comparison not with the immediately preceding model that does not contain the interaction, but rather with the model that does not contain the term at all.

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p
Refit 11	**logitPred** by **speaker**				7150		
Refit 12c	**logitPred**	-0.04	0.02	-2.11	7152 (+2)	0.20	0.66
Refit 13c	**vowel:logitPred**				7138 (-12)	16.31	< 0.001
	-o	0.07	0.02	4.02			

Table 4.6: Summary of the effect of adding **logitPred** to the trimmed model containing a by-speaker random slope for **logitPred**. Becausing adding **logitPred** by itself does not significantly improve model fit (line Refit 12c), its improvement when interacting with **vowel** (Line Refit 13c) is evaluated against the control model (line Refit 11).

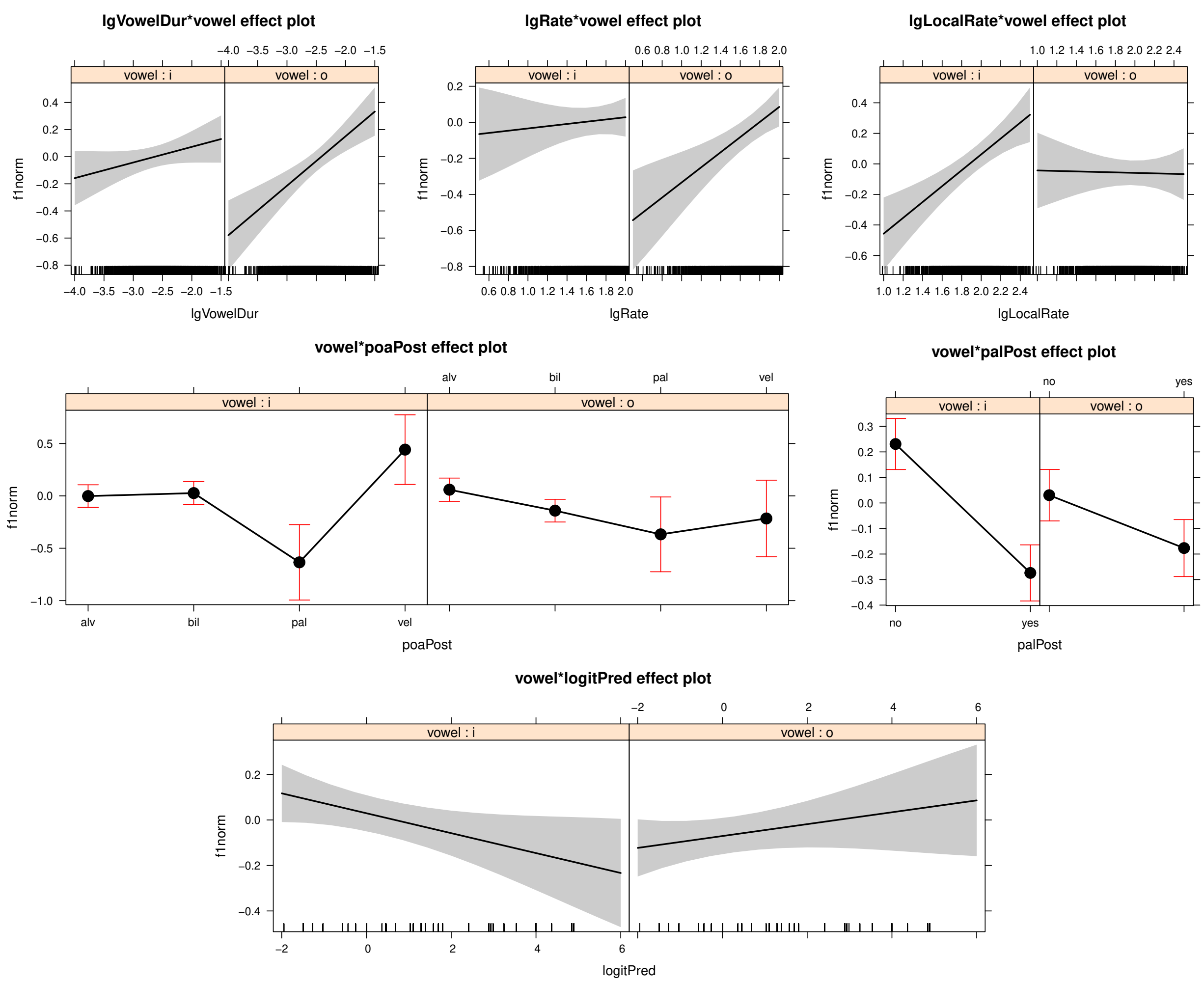


Figure 4.5: Partial effects plot of the F1 model, showing the significant effect of contextual probability (**logitPred**, bottom right panel).

(Nakagawa and Schielzeth, 2013). The interaction between **logitPred** and **vowel** adds only 0.005 to the marginal R^2. In other words, although including this interaction does signficantly improve model fit, it explains only just over half a percent of the variance in F1.

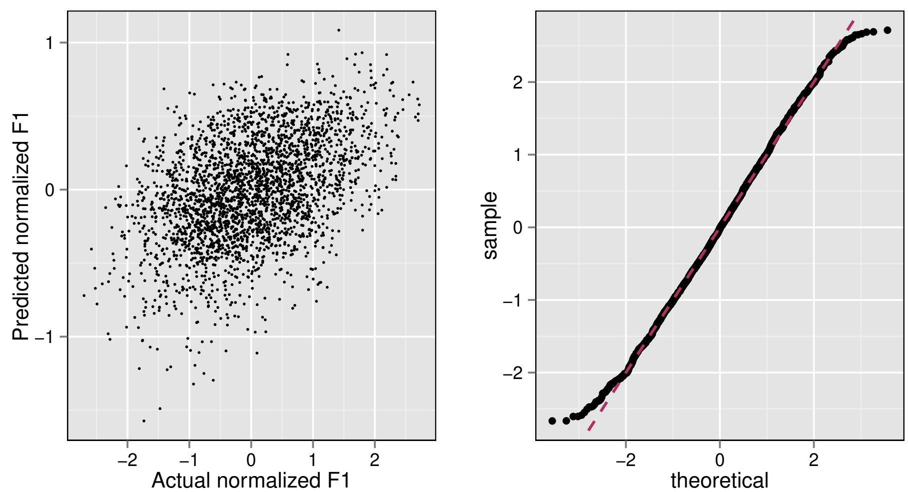


Figure 4.6: Model evaluation plots of the F1 model. Left panel shows the poor model fit, and the right panel shows the non-normality of the residuals.

4.3.3 F2

4.3.3.1 Model summary

The analysis of F2 proceeded exactly in parallel to the model for F1. A summary of this model is given in Table 4.7, and a partial effects plot is shown in Figure 4.7. The prosodic properties that affected F2 were the order of the subject and verb (**order**) and the log-transformed duration of the vowel (**lgVowelDur**). **Order** did not interact with **vowel**; its effect was a uniform increase in F2 when the subject followed the verb. By contrast, **lgVowelDur** showed a strong interaction with **vowel**, such that longer -*i* had a higher F2, while longer -*o* had a lower F2.

Once again, phonetic predictors included the place of articulation of the following consonant (**poaPost**), as well as its palatalization (**palPost**). Unlike its effect on F1, **poaPost** had a stronger effect on -*o* than on -*i*. In particular, a following consonant with a palatal place of articulation increased F2 for -*o*, while following bilabials and velars decreased it. The -*i* suffix was relatively insensitive to the place of articulation of the following consonant. A similar pattern could be observed for the palatalization contrast in following consonants (**palPost**). While F2 increased both for -*i* and -*o* when the following consonant was palatalized, the effect was much stronger for -*o*.

The probabilistic predicters that affected F2 were log-transformed verb frequency (**lgFqFull**) and paradigmatic probability (**lgRatioFull**). Only -*o* showed an effect of **lgVerbFq**, such that F2 increased with higher verb frequency (line 10 in Table 4.7). Both *-o* and *-i* showed an effect of **lgRatioFull**, however (lines 11a-12a). As paradigmatic probability increased, F2 increased for *-i* ($\beta = 0.03, SE(\beta) = 0.01, t = 2.09$), but decreased for *-o* ($\beta = 0.07, SE(\beta) = 0.02, t = -3.53$). Contextual probability (**logitPred**, lines 11b-12b) did not improve the model fit, either as a simple effect ($\chi^2 = 0.02, p = 0.89$) or in interaction with **vowel** ($\chi^2 = 0.58, p = 0.45$). Likewise, the effect of inflectional entropy (**entropyFull**, lines 11c-12c) did not reach significance, either alone ($\chi^2 = 0.05, p = 0.83$) or in interaction ($\chi^2 = 0.97, p = 0.32$).

The model containing **lgRatioFull** in interaction with **vowel** (line 12a in Table 4.7) was selected for further evaluation to determine whether the effect of **lgRatioFull** would still appear after model refinement. Only one random slope improved the fit of the model: **vowel**, grouped by **sentence**. After adding this slope to the model, observations with scaled residuals more than 2.5 standard deviations from 0 were removed. Next, the model containing only control predictors (line 10 in Table 4.7) and the random slope of **vowel** by **sentence** was fit to the trimmed data, yielding the model in line Refit 10 in Table 4.8. Finally, **lgRatioFull** and its interaction with **vowel** were added to this trimmed model, yielding the bottom two lines in Table 4.8. The improvement in model fit from this effect was smaller in this trimmed model than it was in the initial model, but it still reached significance ($\chi^2 = 6.22, p < 0.05$). It is this model whose partial effects are given in Figure 4.7. As the bottom plot shows, F2 for *-i* increased with increasing relative frequency of the plural form compared to the singular, while F2 for *-o* decreased with increasing relative frequency of the singular form compared to the plural. In other words, as the paradigmatic probability of the particular form increased, the articulation of *-i* moved forward, while the articulation of i-o moved backward.

4.3.3.2 Model evaluation

The fit of the F2 model was better than the fit of the F1 model. Figure 4.8 shows the by-now familiar model evaluation plots. The plot on the left shows the normalized F2 values plotted as a function of the values predicted by the model shown in the bottom row of Table 4.8. The spread is smaller than the spread in the equivalent plot for F1 in Figure 4.6, and relatively constant across all values of F2. The qq-plot on the right, however, indicates that the distribution of the residuals deviates from normality in the lower portion of the F2 range. This deviation is enough for a Shapiro-Wilk test for normality to return a signficant result ($W = 0.9963, p < 0.001$).

The marginal R^2 of the trimmed model (line Refit 12a in Table 4.8) is 0.352, indicating that just over a third of the variance in F2 is explained by the fixed effects of the model (Nakagawa and Schielzeth, 2013). This is 0.0046 more than the marginal R^2 of the trimmed control model (line Refit 10 in Table 4.8). In other words, the inclusion of **lgRatio** and its interaction with **vowel** explain about 0.46% of the variance in F2, or about 1.3% of all variance explained by the fixed effects. The conditional R^2 of the full model is 0.429, indicating that less than half of the variance in F2 is explained by fixed and random effects together. Although the fit of this model is better than the fit of the model predicting F1, it is not as good as the duration model.

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p
0	**Intercept**	1.30	0.18	7.18	7358		
1	**vowel**				7360 (2)	0.02	0.88
	-o	-2.92	0.24	-12.32			
2	**lgVowelDur**	0.54	0.06	9.65	7355 (-3)	7.18	< 0.05
3	**vowel:lgVowelDur**				7249 (-106)	107.46	< 0.001
	-o	-1.00	0.09	-11.75			
4	**poaPost**				7227 (-22)	28.46	< 0.001
	bilabial	-0.23	0.07	-3.39			
	palatal	-0.03	0.18	-0.19			
	velar	-0.16	0.15	-1.06			
5	**vowel:poaPost**				7155 (-72)	77.63	< 0.001
	-o, bilabial	-0.39	0.07	-5.69			
	-o, palatal	0.21	0.17	1.20			
	-o, velar	-0.77	0.17	-4.65			
6	**palPost**				7090 (-65)	67.37	< 0.001
	yes	0.42	0.07	6.12			
7	**vowel:palPost**				7061 (-29)	30.11	< 0.001
	-o, yes	0.39	0.07	5.88			
8	**order**				7056 (-5)	7.84	< 0.01
	VS	0.15	0.06	2.38			
9	**lgFqFull**	-0.01	0.02	-0.37	7053 (-3)	4.23	< 0.05
10	**vowel:lgFqFull**				7049 (-4)	6.78	< 0.01
	-o	0.06	0.02	3.79			
11a	**lgRatio**	0.03	0.01	2.09	7051 (+2)	0.01	0.93
12a	**vowel:lgRatio**				7040 (-9)	12.75	< 0.01
	-o	-0.07	0.02	-3.53			
11b	**entropyFull**	-0.03	0.09	-0.36	7051 (+2)	0.05	0.83
12b	**vowel:entropyFull**				7052 (+1)	0.97	0.32
	-o	0.09	0.09	0.99			
11c	**logitPred**	-0.004	0.017	-0.24	7051 (+2)	0.02	0.89
12c	**vowel:logitPred**				7052 (+1)	0.58	0.45
	-o	0.01	0.02	0.76			

Table 4.7: Summary of the full F2 model. The predictors **lgVowelDur** and **lgRatioFull** did not significantly improve the model except in interaction with **vowel**. For this reason, the improvement in model fit contributed by these interaction terms is evaluated in comparison not with the immediately preceding model that does not contain the interaction, but rather with the model that does not contain the term at all.

Model	Predictor	β	$SE(\beta)$	t	AIC (ΔAIC)	χ^2	p
Refit 10	**vowel** by **sentence**				5830		
Refit 11a	**lgRatio**	0.03	0.02	1.58	5832 (+2)	0.00	0.96
Refit 12a	**vowel:lgRatio**				5828 (-2)	6.23	< 0.05
	-o	-0.06	0.02	-2.53			

Table 4.8: Summary of the effect of adding **lgRatio** to the trimmed model containing a by-sentence random slope for **vowel**. Becausing adding **lgRatio** by itself does not significantly improve model fit (line Refit 11a), its improvement when interacting with **vowel** (Line Refit 12a) is evaluated against the control model (line Refit 10).

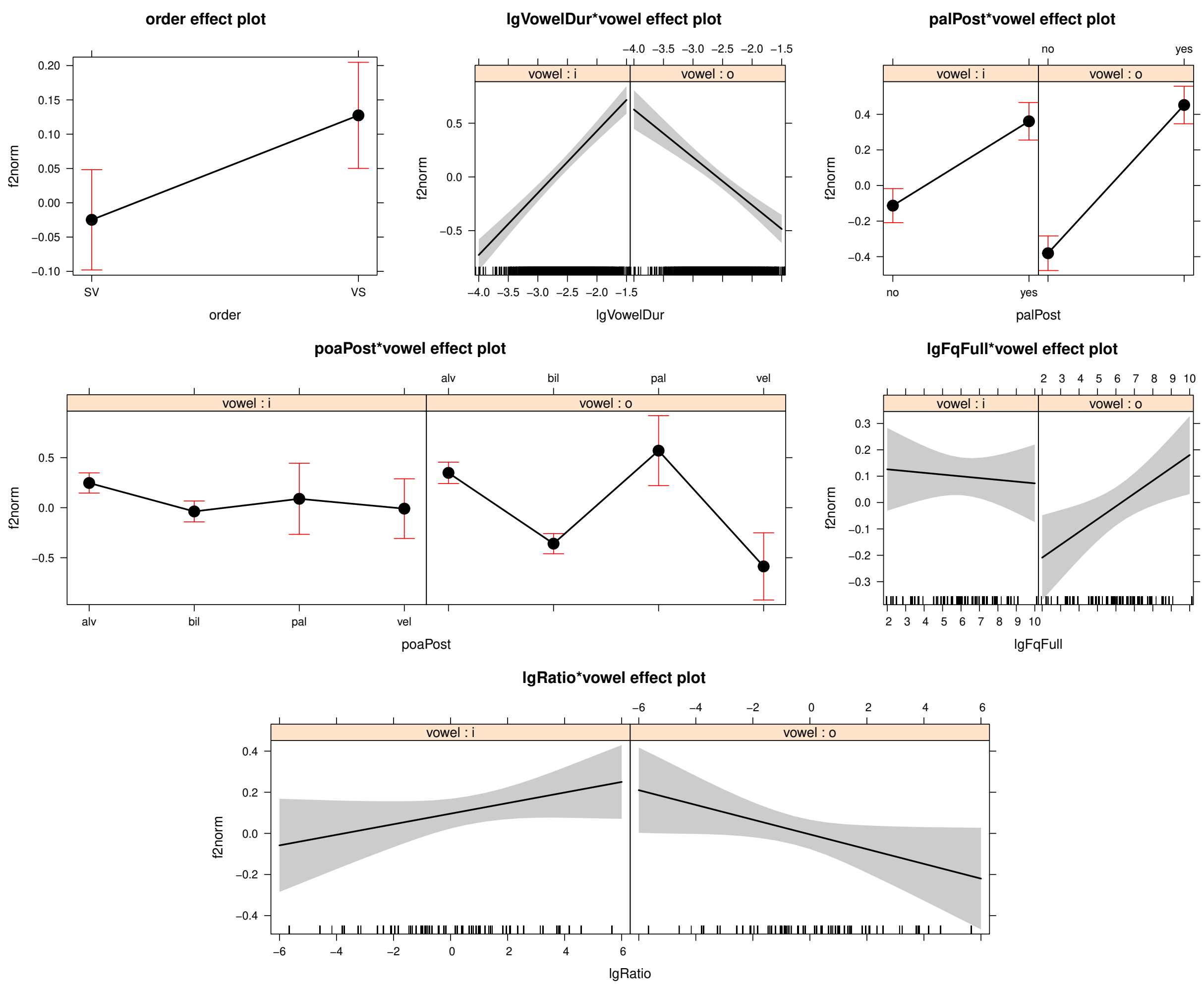


Figure 4.7: Partial effects plot of the F2 model, showing the significant effect of paradigmatic probability (**lgRatioFull**, bottom panel).

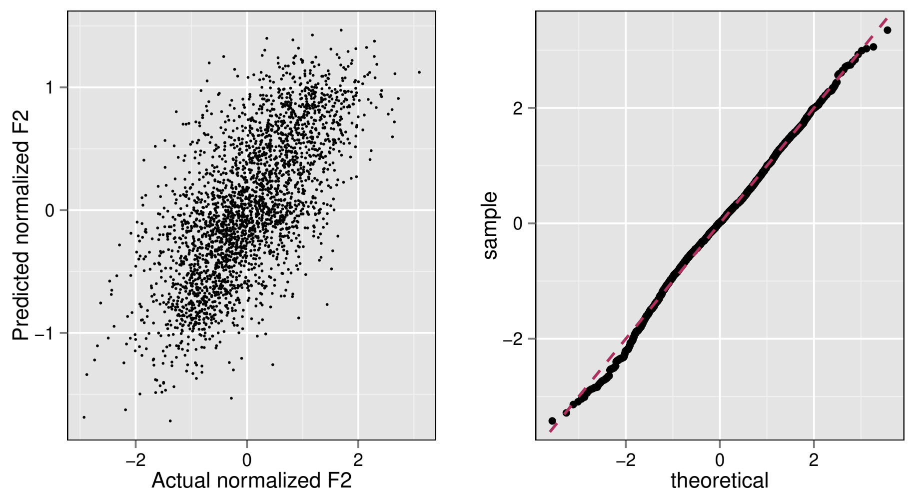


Figure 4.8: Model evaluation plots of the F2 model. Left panel shows the model fit, and the right panel shows the non-normality of the residuals.

4.4 Discussion

The current study examined probabilistic phonetic variation of Russian subject-verb agreement suffixes, with a particular focus on which phonetic features exhibit this variation. The results were largely in line with the findings from Chapter 2. First, as in English, there was no effect of inflectional entropy on the pronunciation of either agreement suffix, despite the frequent finding that entropy does affect retrieval time (Baayen et al., 2006; Bien et al., 2011; Milin et al., 2009; Moscoso del Prado Martín et al., 2004; Tabak et al., 2005). In Chapter 2, I proposed that such an effect might be due to differences in task demands, or simply the result of the fact that English has a simple verbal inflection system. Russian's inflectional system, however, is much richer, and yet no effect of inflectional entropy emerged in this experiment. It is therefore possible that pronunciation variation is not affected by entropy in the same way that lexical retrieval speed is, regardless of how rich the language's inflectional system might be.

Second, effects of both contextual and paradigmatic probability could be observed in the suffix pronunciation. Unlike English, however, these effects emerged not in duration variation, but rather in vowel quality. Contextually probable agreement suffixes showed a consistent decrease in F1, an effect that was larger for singular *-o* than for plural *-i*. For singular *-o*, this is a straightforward type of phonetic reduction. As singular agreement became more probable in a sentence, the suffix that encoded it, *-o*, showed a reduction in F1. This pattern had the effect of reducing the height contrast between mid *-o* and its plural counterpart, high *-i*. The effect on plural *-i* was smaller and more fragile, only emerging after the data had been trimmed to remove outliers. It goes in the opposite direction, however: As plural agreement becomes more probable in the sentence, plural *-i* became higher — enhancing the height of a vowel that is already maximally high. If this effect is replicable, it would indicate that in some very restricted cases, an increase in contextual probability can yield enhancement, rather than reduction. However, the size of this effect is so small, the effect itself is so fragile, and the fit of the F1 model is so generally poor, that it is premature to accept it, especially in the face of so much existing research associating contextual probability with reduction (e.g., Aylett and Turk, 2004, 2006; Bell et al., 2009, 2003; Jurafsky et al., 2001).

Paradigmatically probable agreement suffixes showed a robust phonetic enhancement effect in the form of peripheralization in F2 for both for *-i* and *-o*. As [i] became more paradigmatically probable, F2 increased, thus fronting an already front vowel. As [o] became more paradigmatically probable, F2 decreased, thus backing an already back vowel.

These findings are particularly striking because they are restricted to specific phonetic features. Existing approaches to contextual probabilistic pronunciation variation, such as the Probabilistic Reduction Hypothesis (Jurafsky et al., 2001) or Uniform Information Density (Jaeger, 2010), hold that linguistic units that are more probable in context carry less phonetic detail. These accounts do not, however, make any claims regarding the specific phonetic features that show reduction. Similarly, existing approaches addressing phonetic enhancement of paradigmatically probable morphemes, such as the Paradigmatic Signal Enhancement Hypothesis (Kuperman et al., 2007; Schuppler et al., 2012), hold that morphemes which enjoy greater paradigmatic support also carry more acoustic information, yet they do not make claims regarding the manifestation of that acoustic boost. The Contrast Dependent Pronunciation Variation (CDPV) hypothesis, as outlined in the Introduction of this chapter, falls in line with these previous accounts, while further refining their predictions. Under this hypothesis, probabilistic pronunciation variation affects most strongly those features that define a particularly salient contrast. The results presented here, which show probabilistic pronunciation variation in vowel height and backness, but not in duration, provide evidence in favor of this hypothesis.

What does it mean for a contrast to be "particularly salient"? In the current experiment, relevant contrasts are morphosyntactically defined. The sentence context, with the quantified subject noun phrase, restricts possible verb agreement suffixes to third-person neuter singular or third-person plural. The form of the verb stem, which includes the past tense *-l* marker, restricts these possible third-person suffixes further to the past tense neuter singular *-o*, or the past tense plural *-i*. These contextual properties therefore are responsible for shrinking the verb's inflectional paradigm from 67 forms down to two, and the contrast between these two forms is most salient in the domain of vowel quality. Although *-i* and *-o* do differ in duration, such that *-i* is shorter than *-o*, that is a natural consequence of the fact that [i] is a higher vowel than [o] (Lehiste, 1976). Thus duration is not a salient contrast, because it follows directly from the more salient contrast of vowel height (high [i] vs. mid [o]). This contrast in height, combined with a further contrast in backness (front [i] vs. back [o]) thus determines the domain of variation: Reduction of contextually probable suffixes showed up in F1, which reflects height, while enhancement of paradigmatically probable suffixes showed up in F2, which reflects backness.

4.4.1 Paradigmatic probability and phonetic enhancement

The current chapter found phonetic enhancement of suffixes with high paradigmatic probability. These results are in line with the findings of Kuperman et al. (2007) and Schuppler et al. (2012) in

Dutch, and further shows that the Paradigmatic Signal Enhancement Hypothesis applies to Russian. Yet other work has found an association between higher paradigmatic probability and phonetic *reduction*. Key among these is the work of Hanique and colleagues (Hanique and Ernestus, 2011; Hanique et al., 2010), who examined the pronunciation of the prefix *ge-* and suffix *-t* in Dutch past participles. They found that, as the past participle became more probable within the inflectional paradigm of the verb, *ge-* and *-t* were shortened. How is it that increased paradigmatic probability is sometimes associated with phonetic reduction, and sometimes with enhancement?

Viewed through the lens of CDPV, it is perhaps now possible to account for the contradictory findings. Consider, first, the fact that Hanique and colleagues were examining inflectional forms of verbs. Inflectional paradigms, recall, are sensitive to morphosyntactic context, which can restrict the set of contextually relevant forms to a particular subset. Hanique and colleagues did not consider this subset in their analysis: When examining *-t* they looked at the relative frequency of the past participle to all inflectional forms, and when examining *ge-* they looked at the relative frequency of the past participle to the stem, which is also the first person singular form of the verb. Without knowing the morphosyntactic context in which these forms were used, it is not possible to determine what the contextually determined relevant subpart of the paradigm is. This is key because CDPV applies specifically to enhancement or reduction of *contrasts*, not segments. Consider, for example, the adjectival use of a past participle. When a past participle is used attributively in prenominal position, it has a suffix [-ə] in all cases except before indefinite singular neuter nouns (Donaldson, 2008). If a sufficiently large number of the past participles analyzed by Hanique and colleagues were used attributively, then for a great deal of their data, the morphosyntactic context would restrict the inflectional paradigm to the set of competitors with and without the [-ə] suffix. To enhance this contrast, as the form with the schwa becomes more paradigmatically probable, the schwa would need to be lengthened. On the other side, however, as the paradigmatic probability of the form without the schwa increased, then the remaining word would be shortened. In this way the contextually relevant contrast is enhanced, while at the same time the nature of pronunciation variation of the unsuffixed form resembles reduction.

4.4.2 Methodological concerns

The results of the present study are suggestive, but they are subtle, and the size of the probabilistic effects is quite small. Further, they are open to question due to a number of methodological issues in the experimental design. First, the counterbalancing of stimulus lists means that each participant was exposed to a large number of sentences in which the verb agreement was extremely low-probability. Each list contained 7 sentences that had a plural verb, even though the probability of plural agreement was less than 0.2. There were a further 11 sentences containing a singular verb in each list, even though the probability of plural agreement was greater than 0.8 (see Table 4.2a). Although the nature of the agreement variation with quantified subject noun phrases is often considered stylistic (Timberlake, 2004), sentences at the far ends of the probability spectrum will nevertheless seem ungrammatical if they do not have the most probable form. One particularly opinionated participant would frequently comment after each one of these sentences, *Očen' ploxaja frasa*, 'very bad sentence,' and specified that the problem lay specifically in the verb.

The fact that this participant had sufficient time to make such comments about the sentence reveals the second methodological concern: The fixed time frame alloted to each sentence. In most cases, speakers did not need the full ten seconds to select the appropriate form of the missing word and pronounce the sentence aloud. This meant that they had a few seconds to read back over the stimulus and think about the form of it. Such reflection allowed them to observe patterns in the stimuli that they might not otherwise have noticed, as, indeed, one of them commented after completing the experiment. This time for reflection, combined with the presence of at least 20 sentences in which the agreement sounded unnatural in context, would have highlighted the pattern that many sentences contained quantified subject noun phrases with the "wrong" agreement on the verb. This pattern would have been further evident because none of the fillers contained any systematic grammatical errors. Thus, despite the distraction task of selecting the correct form of a word unrelated to the verb, few speakers were unaware of the focus of the study by the time they had finished the experiment.

If the results reported here could be replicated in an experiment that does not contain these issues, it would lend credence to the conclusion that the patterns of pronunciation variation truly reflect some property of speech production, rather than an artifact of a speaker who has guessed the goal of an unnatural experimental task. Such an experiment is not easily designed, however. It is not feasible to give only the sentences with high-probability verb forms, because one goal is to explore how pronunciation changes across the probability spectrum. The systematicity of the apparent errors could be mitigated, though, by included spurious systematic errors in some of the fillers. Speakers could further be prevented from reflecting on what they have said by providing them with the option to continue to the next trial as soon as they have completed the previous one. This design would still be far from natural — not least because speakers would be asked to produce a large number of ungrammatical sentences — but it would avoid highlighting the nature of the key manipulation in the stimuli.

4.5 Conclusion

Work in probabilistic pronunciation variation has shown that we have a sophisticated understanding of usage patterns in language. We are able to track how often linguistic forms appear in different sentence contexts, and how often different forms of a given lexeme are used. Our pronunciation reflects these patterns in distinct ways. Existing accounts have focused on describing phonetic variation as an absolute system, with certain patterns counting as "reduction," and others "enhancement." This chapter provides a more nuanced view. Reduction and enhancement need not apply to a particular linguistic unit, such as a syllable or a word, but rather to a relationship between linguistic units. This is the basis of CDPV: Pronunciation variation does not reduce or enhance a particular linguistic unit, but a contrast between different possible realizations of that unit. We do not simply adjust phonetic detail across the board. Rather, we adjust the *right* phonetic detail determined by a constellation of contextual and paradigmatic patterns that characterize every utterance we produce.

Chapter 5

Conclusion

5.1 Summary of findings

In this dissertation I have examined how linguistic structure interacts with usage patterns, using pronunciation variation as my main investigative tool. I focused on agreement suffixes, because the probability of using a particular suffix — a usage pattern — can vary with respect to two different structural properties: The agreement relationship with the subject, and the inflectional paradigm of the verb. The results of three experiments have led me to propose the Contrast Dependent Pronunciation Variation hypothesis (CDPV). According to this hypothesis, structure and usage combine to restrict the types of probabilistic pronunciation variation that a speaker employs. By CDPV, pronunciation is not simply "reduced' when certain forms are contextually probable (Jurafsky et al., 2001) or "enhanced" when certain forms are paradigmatically probable (Kuperman et al., 2007). Rather, the phonetic features that vary with respect to contextual or paradigmatic probability are exactly those features which encode salient contrasts between competing forms. Phonetic "reduction" and "enhancement" are not general processes that weaken or strengthen the articulation in predictable, universal ways. Rather, they are targeted adjustments — reductions and enhancement — of the contrasts themselves, and are therefore sensitive to language-specific and perhaps even construction-specific properties.

In the current work, the competing forms have been inflected verb forms in English (Chapter 2) and Russian (Chapter 4). They are produced in contexts that allow variation between singular and plural agreement. In English, that context is sentences with agreement attraction constructions, while in Russian that context is sentences with quantified subject noun phrases. In both languages, therefore, a singular form of the verb will have as its primary competitor the plural form of the same verb. This means that the salient phonetic contrast will be the phonetic feature that distinguishes the single form from the plural form. In English, that feature is duration. The contrast between singular and plural verb forms in English is the presence or absence of an *-s* suffix, and the best way to emphasize the presence of that suffix in contrast to its absence is to lengthen it. As a result, probabilistic phonetic variation shows up in the domain of duration adjustments. When the suffix is contextually probable, it is shorter with respect to the verb stem, and when it is paradigmatically probable, it is longer, both absolutely and with respect to the verb stem. Importantly, there was no equivalent effect on spectral center of gravity. Despite previous work associating duration variation with center of gravity variation (van Son and Pols, 2003; van Son and van Santen, 2005; van Son et al., 2004), center of gravity is not a salient feature that distinguishes the presence of an [s] from its absence. As a result, it did not vary according to the probability of observing that [s]. Only duration showed this type of pronunciation variation.

In Russian, the competing forms were again singular and plural verbs, this time in the past tense, but the salient contrast distinguishing those forms was different. Rather than distinguishing singular from plural by means of the presence or absence of a suffix, as English does, Russian distinguishes those forms by means of different suffixal vowels — in the sentences used here, *-i* for plural and *-o*. As a result, the salient contrast in these sentences was vowel quality. That is why probabilistic pronunciation variation targeted vowel quality alone, without affecting vowel duration. The effect of higher contextual probability seemed to reduce vowel height contrasts by raising the height of singular *-o* towards plural *-i*. The effect of higher paradigmatic probability enhanced vowel backness contrasts both for singular *-o*, which moved farther back, and for plural *-i*, which moved farther forward.

5.2 Mechanisms of paradigmatic enhancement and featural specificity

The key assumption of the CDPV is that the salient contrasts depend on context. Pronunciation variation will target key features only in those cases where there are, in fact, key features to target that distinguish among primary competitors. In the current study, those primary competitors are verb forms that are possible in the context of the sentence — or, more precisely, verb suffixes that are possible in the context of the verb stem and surrounding sentence. In this section I lay out one mechanism of lexical retrieval that can be modified to explain both the enhancement effect observed with paradigmatically probable affixes, and also the specificity of the phonetic features that participate in the enhancement.

Kuperman et al. (2007)'s Paradigmatic Signal Enhancement Hypothesis proposes that forms which are relatively frequent within their paradigms receive more "support" during articulation, and hence are pronounced with greater acoustic detail. The nature of the "support" is not elaborated, however. One possible mechanism that might account for the enhancement effect is segmental competition between multiple activated forms, of the sort proposed by Baese-Berk and Goldrick (2009). This account was originally proposed to explain a type of phonetic strengthening that seems to arise from connections between phonologically related forms — connections that, crucially, are sensitive to contextual factors. Baese-Berk and Goldrick (2009) examined the pronunciation of word-initial voiceless stops in words like *cod* and *cop*. They found that people pronounced the initial stop with a longer voice onset time if the word had a minimal pair neighbor that differed in initial stop voicing. Thus, *cod* had a longer voice onset time in its initial [k] than *cop*, because *cod* has a minimal pair neighbor *god*, while *cop* has no such neighbor **gop*. Importantly, the voice onset time difference was greater when the minimal pair neighbor was presented alongside the target than when it was not presented. The authors interpret these findings with respect to a feedback model of cascading activation. According to this account, the activation of a target form (e.g., *cod*) will activate its segmental content, [k], [ɑ], and [d], but the segmental activation feeds back up to the full-form representation. This has the effect of strengthening the initial activiation of *cod*, but it will also activate other forms that contain those segments, such as *god* (and also *call, kid*, and other phonological neighbors). In order to overcome the competition from *god*, the activation of *cod* must increase enough that activation of the initial [k] of the target can inhibit the initial [g] of the competitor. This activation "boost" is responsible for the more extreme articulatory realization, and hence the longer VOT for *cod* compared to *cop*: *Cod* must compete with *god*, while *cop* need not compete with the nonexistant **gop*. This account also explains why the VOT difference is greater when *cop* is presented along with the target *cod*. If the competitor is presented with the target, then its activation will be higher than it would be from segmental feedback alone, and so the activation of the target must be that much greater in order to inhibit the competitor. In other words, *god*'s existance is a nuisance to *cod*, and it is more of a nuisance when it is present than when it is absent.

Enhancement that springs from higher paradigmatic probability is consistent with Baese-Berk and Goldrick (2009)'s account of segmental competition, with minor adjustments. Assume, to start, that high-frequency forms are stored with a higher resting activation at the level of the word form (Jescheniak and Levelt, 1994). In both the English and Russian experiments, the target forms shared a great deal of segmental content with the competitor forms. In English, the target word entirely contains the competitor (i.e., a singular verb, such as *looks*, contains the plural form, such as *look*), while in Russian the two forms differ only in the final vowel.[1] Activating one form (e.g., *šli*, 'they walked') will therefore send activation to the segments that it contains ([ʃ], [l], [i]), and feedback from those segments will also send activation to the highly similar competitor (*šlo*, 'it walked'). In order to overcome this competition, the target will therefore need an extra boost of activation, so that the target vowel *-i* can overcome the competing *-o*. The greater enhancement that goes along with higher relative frequency is the result of the additive effects of the higher resting activation and the activation boost that is required to overcome competition. When words which already have a higher resting activation than their competitors — those with a higher relative frequency — receive a jolt of additional activation to overcome the segmental competition, then the total activation they send down to their segments will be higher than the signal from word forms with a lower resting activation. This is the source of the strengthened articulation of paradigmatically probable forms.

Under this account, enhancement of paradigmatically probable forms is another instance of the type of enhancement observed in *cod*, which must compete with *god* during retrieval. The fact that paradigmatically probable forms show phonetic enhancement is due to their competition with their phonologically similar paradigm neighbors. Sentence context plays the same role as Baese-Berk and Goldrick (2009)'s stimulus screen. In the same way that displaying both the competitor and

[1] Strictly speaking, they also differ in the palatalization of the past-tense *-l*, because the following plural *-i* induces palatalization, while the following singular *-o* does not.

the target on the screen will activate both forms and enhance segmental competition, producing sentences that allow multiple inflectional forms — such as agreement attraction constructions in English, or quantified subjects in Russian — will also activate both the singular and plural verb forms. A key difference, however, is the role of paradigmatic relations. Baayen and colleagues have shown that word forms are stored with a vast array of connections to other words that belong to the same morphological paradigm, be it inflectional (Baayen et al., 2006, 2008; Bien et al., 2011) or derivational (Baayen et al., 2006, 2007; Moscoso del Prado Martín et al., 2004). Because of these connections, activation of one form will activate morphologically related forms both by means of reverberating feedback driven by their phonological similarity, and by means of their paradigmatic connections at the level of the word form. This means that enhancement arising from segmental competition with morphologically related forms will be qualitatively different from enhancement arising from competition with merely phonologically related forms.

I propose that the consequence of this type of paradigmatically-driven enhancement is the source of the featural specificity observed in both English and Russian. The enormous set of morphophonological phenomena that can be seen in the world's languages all attest to the fact that morphological paradigms are sensitive to sub-segmental phonetic features. In English, for example, the singular *-s* suffix on verbs assimilates to the voicing of the preceding consonant, and is separated from preceding sibilants by the insertion of [ə]. In Russian, final consonants of second-conjugation verbs undergo a wide variety of mutations in the first-person singular form, ranging from relatively normal palatalization of *t* to *č* (e.g., *vstretit'*, 'to meet'; *vstreču*, 'I meet'), to the somewhat bizarre addition of a palatalized *l'* only after labials (e.g., *l'ubit'*, 'to love'; *l'ubl'u*, 'I love'). This interplay between morphology and sub-segmental phonology could be the source of the featural sensitivity observed in Chapters 2 and 4. When target forms are competing with their morphologically related competitors, the close interconnection between them is what makes it possible for the activation boost to target exactly those features that distinguish the relevant forms, thus resulting in an exaggerated articulation of exactly those features.

CDPV explains more than the results presented here. It also aligns with previous findings regarding pronunciation variation of affixes. Consider first the findings of Kuperman et al. (2007), who analyzed the pronunciation of Dutch interfixes as a function of their paradigmatic probability. They focused on two possible interfixes — *-s-* and *-e(n)-* — and observed that each interfix was longer when it was more likely to be used as an interfix in the paradigm of compounds, as determined by the initial noun. In other words, *-s-* was longer when it was used in compounds whose first noun preferred to be followed by the *-s-* interfix, and the *-e(n)-* interfix was longer when it was used in compounds whose first noun preferred to be followed by *-e(n)-*. This is puzzling, because the key contrast between *-s-* and *-e(n)-* is not duration. The key to solving this puzzle is the fact that there is a third type of "interfix" in Dutch compounds — the null interfix. Understandably, Kuperman et al. (2007) did not analyze pronunciation of this affix, as it is difficult to pronounce something with no segmental content in more than one way. Nevertheless, this particular realization of the compound interfix was included in their calculations of paradigmatic probability. The probability of *-s-*, for example, was calculated as the frequency with which *-s-* was used in a particular compound paradigm divided by the size of the entire paradigm. This means that when *-s-* is highly probable with respect to its competitor *-e(n)-*, it is also probable with respect to its silent competitor -∅-. The same is true with *-e(n)-*. The upshot of this three-way competition is that a paradigmatically probable overt interfix is in the same position as a paradigmatically probable singular agreement suffix in English: They are both competing with a null alternative. This means that in both cases, at least one salient contrast will be in duration, which is why both English agreement suffixes and Dutch interfixes showed lengthening with higher paradigmatic probability.

The segmental competition mechanism that I propose underlies CDPV fits in with a second set of findings regarding the pronunciation of morphemes. This effect, recall, is the phenomenon whereby morphemic segments such as [s] in *laps* tend to be longer than homophonic non-morphemic counterparts, such as the [s] in *lapse* (Losiewicz, 1992; Smith et al., 2012; Walsh and Parker, 1983). Morphologically complex forms must deal with exactly the same competing activation from phonologically related words as their simplex counterparts, but they also must overcome competition from morphologically related words. In the work cited here, all words that were analyzed have morphologically related competitors that lack the particular suffix under investigation. For example, plural *laps* has a morphological relationship to singular *lap* than homophonic *lapse* does not have. This relationship means that the unaffixed form *lap* will pose stiffer competition to *laps* than it will to *lapse*. This additional competition means that *laps* requires a stronger jolt of activation, which will result in a correspondingly stronger articulation of the segment that distinguishes it from its stiffest competitor. This is why morphemic segments, such as the [-s] in *laps*, are longer than homophonic non-morephemic segments, such as the [s] in lapse. These segments are subject to CDPV in the way that non-morphemic segments are not, and an increase in duration is the best way to distinguish these words from their stiffest competitors, which differ most saliently in lacking that particular affix.

5.3 Further predictions of CDPV

This hypothesis makes predictions regarding the phonetic realization of probabilistic pronunciation variation. One such prediction is that the language-specific patterns are not actually language-specific, but utterance-specific (cf Foote and Bock, 2012, for a similar argument regarding the apparent cross-linguistic differences in sensitivity to notional effects in agreement variation). In other words, it is not the case that Russian speakers display probabilistic pronunciation variation solely on the F2 dimensions, while English speakers manipulate primarily duration. Rather, these patterns arise because the particular context of interest in the Russian sentences was one in which the key competitors varied along the F2 dimension. In a different context, one in which the key competitors vary according to the presence or absence of a particular morpheme, we would expect to see durational variation of the same sort observed in English. In fact, the genitive of negation provides exactly such a context.

In Russian, sentence subjects canonically appear in nominative case, while direct objects canonically are accusative. Both, however, can appear in genitive case in the context of sentential negation. The conditioning factors of this phenomenon have been heavily studied (Bailyn, 1997; Brown, 1999; Partee and Borschev, 2002; Partee et al., 2011, among many others). The particular cause for using one case over the other is not of great interest here, except inasmuch as it can help in the construction of stimuli in which genitive is more or less contextually probable. The key property of this construction is the fact that the morphological expression of the cases provides a perfect opportunity to test whether probabilistic pronunciation variation is as context-dependent as CDPV predicts. Consider, first, second-declension feminine nouns. The nominative suffix is *-a*, the accusative suffix is *-u*, and the genitive suffix is *-i*. This means that in sentences in which existential negation makes genitive subjects possible, the salient contrast between nominative and genitive will be in vowel height, or F1. In sentences in which verbal negation makes genitive direct objects possible, the salient contrast between accusative and genitive will now be in vowel backness, or F2. Finally, consider inanimate masculine nouns, which have no ending in nominative and accusative, and the suffix *-a* in genitive. For these nouns, either type of sentence will involve a salient contrast between the presence or absence of a suffix — a contrast which, as we saw with English, seems to implicate phonetic variation in duration. Russian genitive of negation therefore provides an excellent test case, within a single language and using a single type of construction, to test the context dependence that CDPV predicts.

A second testable prediction of CDPV is that pronunciation variation should reduce or enhance a specific contrast, rather than some inherent property of the linguistic unit itself. Reducing or enhancing a contrast is a relative procedure, not an absolute one. This means that changes in pronunciation that in some cases might be enhancement would be considered in other cases reduction. Consider, for example, two contextually determined competitors that differ solely by means of an [s]~[ʃ] contrast. The phonetic feature that distinguishes these fricatives from each other is spectral center of gravity (CoG), which is lower for [ʃ] than for [s]. According to CDPV, it is spectral center of gravity that should be most sensitive to probabilistic features. Traditionally, lower spectral center of gravity is considered a type of phonetic reduction (van Son and Pols, 2003; van Son and van Santen, 2005; van Son et al., 2004). Contextually probable [s], being phonetically reduced, should therefore have a lower CoG, while a paradigmatically probable [s], showing phonetic enhancement, should have a higher CoG. According to the work of van Son and colleagues, the same would be true of [ʃ], because a higher spectral center of gravity corresponds to increased articulatory effort for all segments. Under CDPV, however, the reverse would be true of [ʃ], because CDPV predicts that phonetic reduction specifically reduces contrasts, while phonetic enhancement emphasizes them. To change pronunciation of [ʃ] such that its contrast with [s] is reduced, therefore, CoG must be raised, not lowered. Similarly, to change pronunciation of [ʃ] such that it is less similar to [s], CoG must be lowered, not raised.

The fact that a contrast between forms can be sensitive to the set of contextually determined competitors yields a third testable prediction regarding the difference between forms that are probable within an inflectional paradigm, and those that are probable within a derivational paradigm. Inflectional paradigms — or at least the portions that reflect agreement relations — usually are strongly influenced by morphosyntactic relations within the sentence. It is because of these relationships that the inflectional paradigm can be narrowed down to the contextually relevant competitors, thus allowing probabilistic pronunciation variation to target specific relevant features. Determining which form to select from a derivational paradigm, on the other hand, has no such sensitivity to context. Whereas saying *runs* instead of *run* encodes a different agreement relationship between the subject and verb, saying *governance* instead of *governor* changes only the high level message component of the utterance, without reference to morphosyntax. For this reason, derivational paradigms cannot be as easily narrowed down to competitor forms that vary according to a single phonetic feature. The consequence of this distinction is that CDPV should apply only to forms with higher *inflectional* paradigmatic probability. Phonetic enhancement that springs from higher *derivational* paradigmatic probability, of the sort observed by Kuperman et al. (2007), will be more general, affecting multiple phonetic features.

5.4 Cross-linguistic variation in speech production mechanisms

The fact that CDPV accounts for both Russian and English probabilistic pronunciation variation has implications for accounts of cross-linguistic variation in speech production mechanisms. An open question in research in speech production is the extent to which speakers of different languages employ separate processes during production. Of particular interest here are differences in processing of agreement relations and differences in mechanisms of lexical retrieval.

5.4.1 Differences in agreement processing

Do speakers of different languages employ different strategies when they process agreement? Some work has suggested that agreement mechanisms show cross-linguistic differences that can be attributed to the structure of the morphological system. Franck et al. (2008) found that gender agreement of predicative adjectives varied systematically across French, Italian, and Spanish as a function of the validity of various cues that are used to identify gender both in the adjective and in the controlling noun: “[T]he system responsible for computing agreement is finely tuned to the distribution of gender markers in the language” (pg. 354). Differing rates of agreement attraction similarly suggest that production of agreement varies across languages. Vigliocco et al. (1996b) found greater effects of distributivity in French and Dutch, compared to English and tentatively suggested that this might be due to higher degrees of morphological richness in the former two languages. On the other hand, however, the remarkably low rate of agreement attraction in Russian led Lorimor et al. (2008) to suggest that morphologically complex languages like Russian are *less* affected by notional properties of number than simpler ones, like English.

Yet there is some evidence for the claim that processing systems for agreement relations are similar. In contrast to both Vigliocco et al. (1996b) and Lorimor et al. (2008), Bock et al. (2012) found similar effects of notional number in both morphologically rich Spanish and morphologically poor English. Foote and Bock (2012) proposed that apparent effects of morphological richness on notional number do not spring from differences across languages, but simply from how much morphology a given utterance contains. Eberhard et al. (2005)’s prominent Marking and Morphing model of number agreement was developed to predict usage variability in English, but it has been extended quite accurately to Dutch (Antón-Méndez and Hartsuiker, 2010) and Hebrew (Deutsch and Dank, 2009), and (less accurately) to Serbian (Mirković and MacDonald, 2013).

Although this dissertation was not designed to answer this question, the results do fall on the side of different mechanisms across languages. In English, contextual reduction of agreement morphemes interacted with verb frequency, while in Russian, there was no such interaction. It could be that this difference is the result of language-dependent differences in the processing of agreement relations. In English, I claim that the source of the interaction between contextual probability and word frequency is that high-frequency inflected verb forms are stored as whole words, while low-frequency forms are not. This means that retrieving the appropriately inflected form of the verb in English requires the ability to link a desired feature, such as singular, both to a decomposed representation of the suffix alone, and to a whole-word representation of the inflected verb that happens to bear the matching feature. The absence of such an interaction in Russian could indicate that the process of resolving grammatical agreement is insulated from whole word representations. Another language-dependent reason for this difference could be that it is a simple consequence of the richer inflectional system. If each lexeme can have dozens of different forms, then any given word form will be used far less often, and will therefore have a weaker whole word representation than in a language where the maximum inflectional paradigm has only five members.

Yet it is possible that the differing effects of contextual probability have a more mundane source. In order for morphosyntactic context to restrict an inflectional paradigm to a particularly salient subset, it must be possible for that context to allow a certain degree of variation. In Chapter 2, this context was provided by agreement attraction constructions. Those are in most analyses a type of speech error, so the probability of plural agreement in a sentence with such a construction is, at heart, the probability of producing an ungrammatical utterance (although see Staub, 2009, 2010). The construction allowing variable number agreement in Russian that was exploited here represents something different. Barring sentences in which the particular verb form was very low probability, it is quite possible to use either singular or plural agreement perfectly grammatically. The variation is stylistic, not grammatical. It may therefore be the case that the production process for these two different types of sentences differ, resulting in different effects on pronunciation of the relevant morphemes. This could be the reason that contextual probability interacted with verb frequency in English, but not in Russian. In other words, the differences may not have been due to the language, but simply due to the construction that was used.

5.4.2 Differences in morphological processing

Beyond the question of whether speakers of different languages process agreement relations differently, there is the similarly fraught question of whether speakers of different languages process morphological structure differently. A large body of work on morphological retrieval has suggested that the nature of the connections between derivationally related words is language dependent. For example, in non-masked priming studies, even related languages like English and German show that morphologically related words have qualitatively different relationships in the mental lexicon. English words will prime morphologically related counterparts only if there is a transparent semantic relationship between the two (Marslen-Wilson et al., 1994; Rastle et al., 2000), while German words do not require such a relationship (Smolka et al., 2014). This language-dependence has even been observed in the brains of bilinguals: MRI experiments showed that Hebrew-English bilinguals make use of the semantic link between morphologically related words when using English, but not when using Hebrew (Bick et al., 2011).

It is not yet clear whether Russian patterns like English, showing semantics-dependent morphological priming, or like German, showing semantics-blind priming. Existing work on morphological priming in Russian has been confined to the masked priming paradigm (Kazanina, 2011; Kazanina et al., 2008), which consistently shows semantic-independent priming regardless of the target language (Diependaele et al., 2005; Longtin et al., 2003; Rastle et al., 2004). Yet the structure of the mental lexicon determines the nature of the link between paradigmatic probability and pronunciation variation. If a language has links between semantically opaque forms in a derivational paradigm, this will have consequences for the paradigmatic probability of a given word. If CDPV is accurate, then the fact that it accounts for both English and Russian data presented here suggests that these two languages share similar lexical retrieval mechanisms.

5.5 Conclusion

The contribution of this dissertation is threefold. First, more methodological than theoretical, I have presented a multivariate statistical model that can be used to predict the probability of observing plural agreement in a Russian sentences with quantified subject noun phrases. Although the variation between singular and plural agreement in these sentences has been the topic of some heavy study, until now there has been very little in the way of multivariate analysis. Chapter 3 of this dissertation presents exactly such an analysis. The results confirm that previous, univariate approaches to this phenomenon have not fallen prey to the types of pitfalls associated with such approaches, and the interactions that emerged in the model provide a little more nuance to the question of which factors affect agreement variation with quantified subjects in Russian.

The second contribution of this dissertation is an extension of the research on the effect of contextual probability on pronunciation. Whereas previous work has observed that higher contextual probability can condition phonetic reduction on segments, syllables, words, and multiword collocations, I focus on morphemes. The effects in Russian show a reduction of the height contrast when the singular suffix is more probable, but the effect on the plural suffix is more fragile, and not entirely reliable. In English, increasing contextual probability results in a decrease in suffix duration relative to the stem for low-frequency lexemes, but the effect is reversed for high-frequency lexemes. Importantly, this effect is driven entirely by stem duration. I argue that the effect falls entirely in line with previous work, under the assumption that low-frequency forms are retrieved in a decomposed manner, while high-frequency forms are retrieved via whole-word representations.

The final, and primary contribution is an extension of recent observations that higher paradigmatic probability, unlike higher contextual probability, seems to be associated with phonetic enhancement of the affixes in question. Previous work has observed this effect on Dutch interfixes (Kuperman et al., 2007) and past tense suffixes (Schuppler et al., 2012), and on English plural noun suffixes (Hay et al., 2012). I show that the same pattern can be observed in subject-verb agreement suffixes in English and Russian. I further propose, in a hypothesis that I call Contrast Dependent

Pronunciation Variation (CDPV), that the type of phonetic variation that appears is determined by the context of the sentence and the morphophonological expression of the possible forms that can occur in that context. Specifically, paradigmatically probable forms will show phonetic enhancement on the features that distinguish them from contextually licensed competitors. I propose that the source of this enhancement effect — and, indeed, other sources of morphological phonetic enhancement (Losiewicz, 1992; Smith et al., 2012; Walsh and Parker, 1983) — is the result of segmental competition of the sort proposed by Baese-Berk and Goldrick (2009), which targets subsegmental features by means of paradigmatic connections in the mental lexicon.

The original goal of this dissertation was to explore the intersection of structure and usage during speech production, at both the sentence level and the word level. The contextual probability of observing a given agreement morpheme is, at its core, the sentence-level union of structure (subject-verb agreement) with usage (probability of using a particular linguistic form). The paradigmatic probability of using a given inflectional form is, at its core, the word-level union of structure (the verb's inflectional paradigm) with usage (the probability of selecting a form from that paradigm). CDPV describes how both of these unions affect pronunciation, and further pulls in a third intersection between the sentence level and the word level. Sentence-level structure is key in reducing the word-level paradigm to the most salient subset, and it is that subset that determines how paradigmatic probability affects pronunciation. Broadly speaking, CDPV describes one corner of a vastly complex speech production system, and provides an understanding of how people navigate the sea of linguistic patterns that surround them.

References

Alegre, M. and Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40:41 – 61.

Antón-Méndez, I. and Hartsuiker, R. J. (2010). Morphophonological and conceptual effects on Dutch subject-verb agreeement. *Language and Cognitive Processes*, 25(5):728–748.

Antón-Méndez, I., Nicol, J. L., and Garrett, M. F. (2002). The relationship between gender and number agreement processing. *Synax*, 5(1):1–25.

Aylett, M. and Turk, A. (2004). The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence, and Duration in Spontaneous Speech. *Language and Speech*, 47(1):31–56.

Aylett, M. and Turk, A. (2006). Language redundancy predicts syllabic duration and spectral characteristics of vocalic syllable nuclei. *Journal of the Acoustic Society of America*, 119:3048–3058.

Baayen, H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge University Press, Cambridge.

Baayen, H., Piepenbrock, R., and Gulikers, L. (1993). The CELEX Lexical Database. CD-ROM.

Baayen, R. H. (2009). Corpus linguistics in morphology: morphological productivity. In Lüdeling, A. and Kytö, M., editors, *Corpus linguistics: an international handbook*, pages 900–919. Mouton De Gruyter, Berlin.

Baayen, R. H., Dijkstra, T., and Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37:94 – 117.

Baayen, R. H., Feldman, L. B., and Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 55:290 – 313.

Baayen, R. H., Levelt, W. M. J., Schreuder, R., and Ernestus, M. (2008). Paradigmatic structure in speech production. *Chicago Linguistics Society*, 43:1–29.

Baayen, R. H. and Moscoso del Prado Martín, F. (2005). Semantic density and past-tense formation in three Germanic Languages. *Language*, 81:666 – 698.

Baayen, R. H. and Schreuder, R. (1999). War and peace: Morphemes and full forms in a noninteractive activation parallel dual-route model. *Brain and Language*, 68:27 – 32.

Baayen, R. H., Wurm, L., and Aycock, J. (2007). Lexical dynamics for low-frequency complex words. *The Mental Lexicon*, 2(3):419 – 463.

Badecker, W. and Kuminiak, F. (2007). Morphology, agreement, and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language*, 56:65–85.

Baese-Berk, M. and Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes*, 24(4):527–554.

Bailyn, J. F. (1997). Genitive of Negation is Obligatory. In Browne, W., editor, *Formal Approaches to Slavic Linguistics #4: The Cornell Meeting*, pages 84 – 114.

Baker, M. C. (2008). *The Syntax of Agreement and Concord*. Cambridge University Press, Cambridge, UK.

Barlow, M. and Kemmer, S., editors (2000). *Usage based models of language*. CSLI Publications, Stanford, CA.

Barnes, J. (2006). *Strength and weakness at the interface: position neutralization in phonetics and phonology*. Mouton De Gruyter.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3):255–278.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2013). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.0-5.

Baugh, J. (1980). A reexamination of the Black English copula. In Labov, W., editor, *Locating language in time and space*, Quantitative analyses of linguistic structure, pages 83–106. Academic Press, New York.

Beckman, M. and Edwards, J. (1990). Lengthenings and shortenings and the nature of prosodic constituency. In Kingston, J. and Beckman, M., editors, *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, pages 152 – 178. Cambridge University Press.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., and Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60:92–111.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., and Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustic Society of America*, 113(2):1001–1024.

Bertram, R., Schreuder, R., and Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, afaffix homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26(2):489 – 511.

Bick, A., Goelman, G., and Frost, R. (2011). Hebrew brain vs. English brain: Language modulates the way it is processed. *Journal of Cognitive Neuroscience*, 23(9):2280 – 2290.

Bien, H., Baayen, R. H., and Levelt, W. J. M. (2011). Frequency effects in the production of Dutch deverbal adjectives and inflected verbs. *Language and Cognitive Processes*, 26(4-6):683–715.

Bock, J. K. and Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, 8:57–99.

Bock, J. K., Eberhard, K. M., Cutting, J. C., Meyer, A. S., and Schriefers, H. (2001). Some attractions of verb agreement. *Cognitive Psychology*, 43:83 – 2001.

Bock, K., Butterfield, S., Cutler, A., Cutting, J. C., Eberhard, K. M., and Humphreys, K. R. (2006). Number agreement in British and American English: Disagreeing to agree collectively. *Lan-*

guage, 82(1):64–113.

Bock, K., Carreiras, M., and Meseguer, E. (2012). Number meaning and number grammar in English and Spanish. *Journal of Memory and Language*, 66:17–37.

Bock, K. and Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31(1):99 – 127.

Bock, K. and Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23(1):45 – 93.

Bod, R., Hay, J., and Jannedy, S., editors (2003). *Probabilistic Linguistics*. The MIT Press.

Boersma, P. and Weenink, D. (2009). Praat: doing phonetics by computer. Version 5.1, retrieved 31 January 2009 from `http://www.praat.org`.

Bresnan, J., Cueni, A., Nikitina, T., and Baayen, R. H. (2007). Predicting the dative alternation. In Bouma, G., Krámer, I., and Zwarts, J., editors, *Cognitive foundations of interpretation*. Royal Netherlands Academy of Science, Amsterdam.

Bresnan, J. and Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language*, 86(1):168–213.

Brown, S. (1999). *The Syntax of Negation in Russian*. CSLI.

Brysbaert, M., New, B., and Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, 44(4):991–997.

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford University Press, New York.

Bybee, J. and Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics*, 37:575–596.

Bybee, J. L. and Slobin, D. I. (1982). Rules and schemas in the development and use of the English past Tense. *Language*, 58(2):265 – 289.

Carroll, L. (1871). *Through the looking glass, and what Alice found there*. Macmillan.

Cedergren, H. J. and Sankoff, D. (1974). Variable rules: Performance as a statistical reflection of competence. *Language*, 50(2):333 – 355.

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

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, USA.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. MIT Press.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, 2:225 – 252.

Corbett, G. (1979). Predicate Agreement in Russian. Birmingham Slavonic Monographs, No. 7, University of Birmingham, Department of Russian Language & Literature.

Corbett, G. (1998). Agreement in Slavic. Bloomington, Indiana. Papers presented at the *Workshop on Comparative Slavic Morphosyntax*.

Corbett, G. G. (1983). *Hierarchies, Targets and Controllers: Agreement Patterns in Slavic*. Pennsylvania State University Press, University Park, PA.

Corbett, G. G. (2006). *Agreement*. Cambridge textbooks in linguistics. Cambridge University Press.

Crystal, T. H. and House, A. S. (1988). Segmental durations in connected-speech signal: Current results. *Journal of the Acoustic Society of America*, 83(4):1553 – 1573.

de Saussure, F. (1916/1959). *Course in General Linguistics*. Columbia University Press.

Delvaux, V. and Soquet, A. (2007). The influence of ambient speech on adult speech production through unintentional imitation. *Phonetica*, 64:145–173.

Demberg, V., Sayeed, A. B., Gorinski, P. J., and Engonopoulos, N. (2012). Syntactic surprisal affects spoken word duration in conversational contexts. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 356–367.

Deutsch, A. and Dank, M. (2009). Conflicting cues and competition between notional and grammatical factors in producing number and gender agreement. *Journal of Memory and Language*, 60:112–143.

Diependaele, K., Sandra, D., and Grainger, J. (2005). Masked cross-modal orthographic priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes*, 20(1/2):75 – 114.

Divjak, D. (2008). On (in)frequency and (un)acceptability. In Lewandowka-Tomaszczyk, B., editor, *Corpus Linguistics, Computer Tools, and Applications — State of the Art*, Łódź studies in language, pages 213 – 233. Peter Lang, Frankfurt am Main.

Donaldson, B. (2008). *Dutch : a Comprehensive Grammar*. Routledge, New York, second edition.

Eberhard, K. M., Cutting, J. C., and Bock, K. (2005). Making Syntax of Sense: Number Agreement in Sentence Production. *Psychological Review*, 112(3):531 – 559.

Erker, D. and Guy, G. R. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*, 88(3):526–557.

Fidelholtz, J. L. (1975). Word frequency and vowel reduction in English. In Grossman, R. R., San, L. J., and Vance, T. J., editors, *Papers from the eleventh regional meeting of the Chicago Linguistic Society*. Chicago Linguistic Society.

Flemming, E. (2005). A phonphonetic-based model of phonological vowel reduction. Unpublished manuscript, MIT.

Foote, R. and Bock, K. (2012). The role of morphology in subject-verb number agreement: A comparison of Mexican and Dominican Spanish. *Language and Cognitive Processes*, 27(3):429–461.

Forster, K. I. and Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12:627–635.

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., and Frauenfelder, U. H. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, 23(3):329–374.

Franck, J., Vigliocco, G., and Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, 17(4):371–404.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84(3):474–496.

Gahl, S. and Garnsey, S. M. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, 80(4):748–775.

Gahl, S., Yao, Y., and Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, 66(4):789 – 806.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., and Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37(1):58–93.

Geluykens, R. (1988). On the myth of rising intonation in polar questions. *Journal of Pragmatics*, 12:467 – 485.

Gillespie, M. and Pearlmutter, N. J. (2013). Against structural constraints in subject-verb agreement production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 39(2):515–528.

Gilquin, G. and Gries, S. T. (2009). Corpora and experimental methods: a state-of-the-art review. *Corpus Linguistics and Linguistic Theory*, 5(1):1–26.

Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., and Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. In *Chicago Linguistic Society*, volume 35, pages 151–166. Citeseer.

Gries, S. T. (2003). *Multifactorial analysis in corpus linguistics: A study of particle placement*. Continuum, New York.

Guy, G. R. (1991). Explanations in variable phonology: An exponential model of morphological constraints. *Language Variation and Change*, 3(1):1–22.

Guy, G. R. and Bayley, R. (1995). On the Choice of Relative Pronouns in English. *American Speech*, 70(2):pp. 148–162.

Halle, M. (1959). *The sound pattern of Russian: A linguistic and acousical investigation*. Mouton & Co., The Hague.

Hanique, I. and Ernestus, M. (2011). Final /t/ reduction in Dutch past-participles: The role of word predictability and morphological decomposability. In *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011)*, pages 2849–2852, Florence, Italy.

Hanique, I., Schuppler, B., and Ernestus, M. (2010). Morphological and predictability effects on schwa reduction: The case of Dutch word-initial syllables. In *In Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech 2010)*, pages 933–936, Makuhari, Japan.

Hartsuiker, R. J., Antón-Méndez, I., and van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, 45:546 – 572.

Haskell, T. R. and MacDonald, M. C. (2003). Conflicting cues and competition in subject-verb agreement. *Journal of Memory and Language*, 48(4):760 – 778.

Haskell, T. R., Thornton, R., and MacDonald, M. C. (2010). Experience and grammatical agreement: Statistical learning shapes number agreement production. *Cognition*, 114(2):151 – 164.

Hay, J. (2001). Lexical frequency in morphology: is everthing relative? *Linguistics*, 39(6):1041–1070.

Hay, J. and Baayen, H. (2001). Parsing and Productivity. In Booij, G. and Marle, J. V., editors, *Yearbook of Morphology: 2001*, pages 203–235. Kluwer Academic Publishers, Dordrecht.

Hay, J., Kuperman, V., and Saarinen, P. (2012). Paradigmatic structure in plural duration: A cross dialectal study. Paper presented at the 13th Language and Society Conference in Aukland, New Zealand.

Hay, J. and Schreier, D. (2004). Reversing the trajectory of language change: Subject-verb agreement with *be* in New Zealand English. *Language Variation and Change*, 16:209–235.

Hay, J. B. (2003). *Causes and Consequences of Word Structure*. Routledge, New York and London.

Heerman, V. D. (1930). Animal Crackers [Motion picture]. United States: Paramount-Publix Corp.

Howes, D. (1957). On the relation between the intelligibility and frequency of occurrence of English words. *Journal of the Acoustical Society of America*, 29(2):296 – 305.

Howes, D. H. and Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, 41:401 – 410.

Humphreys, K. and Bock, K. (2005). Notional number agreement in English. *Psychonomic Bulletin & Review*, 12:689–695.

Jaeger, T. F. (2010). Redundancy and reduction: speakers manage syntactic information density. *Cognitive Psychology*, 61:23–61.

Jescheniak, J. and Levelt, W. J. M. (1994). Word Frequency Effects in Speech Production: Retrieval of Syntactic Information and of Phonological Form. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20(4):824–843.

Jurafsky, D. (2002). Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In Bod, R., Hay, J. B., and Jannedy, S., editors, *Probabilistic Linguistics*. MIT Press.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J. and Hopper, P., editors, *Frequency and the emergence of linguistic structure*, Typological studies in language 45, pages 229–255. John Benjamins, Amsteram.

Kazanina, N. (2011). Decomposition of prefixed words in Russian. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 37(6):1371–1390.

Kazanina, N., Dukova-Zheleva, G., Geber, D., Kharlamov, V., and Tonciulescu, K. (2008). Decomposition into multiple morphemes during lexical access: A masked priming study of Russian nouns. *Language and Cognitive Processes*, 23(6):800 – 823.

Kuperman, V. and Bresnan, J. (2012). The effects of construction probability on word durations during spontaneous incremental sentence production. *Journal of Memory and Language*, 66:588–611.

Kuperman, V., Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2007). Morphological predictability and acoustic salience of interfixes in Dutch compounds. *Journal of the Acoustical Society of America*, 121:2261–2271.

Kuvšinskaja, J. M. (2012). Soglasovanie skazujemogo s podležaščim, vyražennym sčetnym oborotom. Problemy russkoj stilistiki po dannym NKRJ. Online at http://studiorum.ruscorpora.ru.

Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, 45(4):pp. 715–762.

Labov, W. (1972). Is the Black English Vernacular a separate system? In *Language in the inner city: Studies in the Black English Vernacular*, pages 36–64. University of Pennsylvania Press, Philadelphia.

Lambrecht, K. and Polinsky, M. (1997). Typological Variation in Sentence-Focus Constructions. In Singer, K., Eggert, R., and Anderson, G., editors, *CLS 33: Papers from the panels on Linguistic Ideologies in Contact; Universal Grammar; Parametres and Typology; The Perception* *of Speech and other Acoustic Signals*, pages 189–206. Chicago Linguistics Society.

Laudanna, A., Badecker, W., and Caramazza, A. (1992). Processing inflectional and derivational morphology. *Journal of Memory and Language*, 31:333 – 348.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *The Journal of the A*, 6(2):2018 – 2024.

Lehiste, I. (1976). Suprasegmental features of speech. In Lass, N., editor, *Contemporary issues in experimenal phonetics*, Perspectives in neurolinguistics and psycholinguistics, pages 225 – 239. Academic Press, New York.

Liberman, A., Cooper, F., Shankweiler, D., and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6):431–461.

Lieberman, P. (1963). Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech. *Language and Speech*, 6(3):172–187.

Longtin, C.-M., Segui, J., and Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, 18(3):213 – 334.

Lorimor, H., Bock, K., Zalkind, E., Sheyman, A., and Beard, R. (2008). Agreement and attraction in Russian. *Language and Cognitive Processes*, 23(6):769–799.

Losiewicz, B. (1992). *The effect of frequency on linguistic morphology*. PhD thesis, University of Texas at Austin.

Luke, S. G. and Christianson, K. (2011). Stem and whole-word frequency effects in the processing of inflected verbs in and out of a sentence context. *Language and Cognitive Processes*, 26(8):1173 – 1192.

Lunden, A. (2010). The weight of final syllables in English. In *Proceedings of the 28th West Coast Conference on Foml Linguistics*.

Lunden, A. (2013). Reanalyzing final consonant extrametricality: A proportional theory of weight. *Journal of Comparative German Linguistics*, 16:1 – 31.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., and Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101(1):3–33.

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavioral Research Methods*, 44(2):314–324.

McGinnies, E., Comer, P. B., and Lacey, O. L. (1952). Visual-recognition thresholds as a function of word length and word frequency. *Journal of Experimental Psychology*, 44(2):65 – 69.

Meyer, A. S. and Bock, J. K. (1999). RRepresentation and processes in the production of pronouns: Some Perspectives from Dutch. *Journal of Memory and Language*, 41:281–301.

Milin, P., Kuperman, V., Kostić, A., and Baayen, R. H. (2009). Paradigms bit by bit: An information-theoretic approach to the processing of paradigmatic structure in inflection and derivation. In Blevins, J. P. and Blevins, J., editors, *Analogy in grammar: form and acquisition*, pages 214 – 252. Oxford University Press, Oxford.

Mirković, J. and MacDonald, M. C. (2013). When singular and plural are both gramatical: Semantic and morphophonological effects in agreement. *Journal of Memory and Language*, 69:277–298.

Moscoso del Prado Martín, F., Kostić, A., and Baayen, R. H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, 94:1 – 18.

Nagel, H. N., Shaprio, L. P., Tuller, B., and Nawy, R. (1996). Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing. *Journal of Psycholinguistic Research*, 25(2):319 – 344.

Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4:133 – 142.

New, B., Brysbaert, M., Segui, J., Ferrand, L., and Rastle, K. (2004). The processing of singular and plural nouns in French and English. *Journal of Memory and Language*, 51:568 –585.

Nichols, J., Rappaport, G., and Timberlake, A. (1980). Subjec, Topic and Control in Russian. In *Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society*, pages 372–386.

Padgett, J. and Tabain, M. (2005). Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica*, 62(1):14–54.

Partee, B. and Borschev, V. (2002). Genitive of negation and scope of negation in Russian existential sentences. In Toman, J., editor, *Formal Approaches to Slavic Linguistics #10: The Second Ann Arbor Meeting 2001*, pages 181 – 200.

Partee, B. H., Borschev, V., Paducheva, E. V., Testelets, Y., and Yanovich, I. (2011). Russian genitive of negation alternations: The role of verb semantics. *Scando-Slavica*, 57(2):135 – 159.

Patton, H. B. (1969). *A study of the agreement of the predicate with a quantitative subject in contemporary Russian*. PhD thesis, University of Pennsylvania.

Pierrehumbert, J. and Hirschberg, J. (1990). The meaning of inintonation contours in the interpretation of discourse. In Cohen, P. R., Morgan, J., and Pollack, M., editors, *Intentions in Communication*. MIT Press.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. HarperCollins, New York.

Pinker, S. and Ullman, M. (2002a). Combination and structure, not gradedness, is the issue: reply to McClelland and Patterson. *TRENDS in Cognitive Sciences*, 6(11):472–474.

Pinker, S. and Ullman, M. (2002b). The past and future of the past tense. *TRENDS in Cognitive Sciences*, 6(11):456–463.

Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, 62:146–159.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., and Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustic Society of America*, 90:2956 – 2970.

Prince, A. and Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. Blackwell.

Purcell, E. T. (1979). Formant frequency patterns in Russian VCV utterances. *Journal of the Acoustical Society of America*, 66(6):1691 – 1702.

R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rastle, K., Davis, M., and New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, 11(6):1090–1098.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., and Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, 15(4-5):507–537.

Riordan, B. (2007). There's more than two ways to say it: Modeling nonprestige *there's*. *Corpus Linguistics and Linguistic Theory*, 3(2):233–279.

Robblee, K. E. (1993a). Individuation and Russian Agreement. *The Slavic and East European Journal*, 37(4):pp. 423–441.

Robblee, K. E. (1993b). Predicate Lexicosemantics and Case Marking under Negation in Russian. *Russian Linguistics*, 17(3):pp. 209–236.

Robblee, K. E. (1997). The interaction of Russian word order, agreement and case marking. In Simon-Vandenbergen, A.-M., Davidse, K., and Noël, K., editors, *Reconnecting Language: Morphology and Syntax in Functional Perspectives*, pages 227–248. John Benjamins.

Santesteban, M., Pickering, M. J., and Branigan, H. P. (2013). The effects of word order on subject-verb and object-verb agreement: Evidence from Basque. *Journal of Memory and Language*, 68:160–179.

Scheibman, J. (2000). I dunno: A usage-based account of the phonological reduction of *don't* in American English conversation. *Journal of Pragmatics*, 32:105–124.

Schuppler, B., van Dommelen, W. A., Koreman, J., and Ernestus, M. (2012). How linguistic and probabilistic properties of a word affect the realization of its final /t/: Studies at the phonemic and sub-phonemic level. *Journal of Phonetics*, 40:595–607.

Shriberg, E. (1999). Phonetic consequencies of speech disfluency. Technical report, Speech Technology and Research Laboratory.

Simonov, K. (1958). *Tovarišči po oružiju [Comrades in Arms]*. Sovietskij pisatel', Moscow.

Smith, R., Baker, R., and Hawkins, S. (2012). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics*, 40:689–705.

Smolka, E., Preller, K., and Eulitz, C. (2014). '*Verstehen*' ('understand') primes '*stehen*' ('stand'): Morphological structure overrides semantic compositionality in the lexical representation of German complex verbs. *Journal of Memory and Language*, 72:16 – 36.

Solomon, E. S. and Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, 49(1):1 – 46.

Sprenger, S. (2003). *Fixed expression and the production of idioms*. PhD thesis, Max Planck Institute for Psycholingustics.

Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, 60(2):308 – 327.

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition*, 114:447 – 454.

Sugahara, M. and Turk, A. (2009). Durational correlates of English sublexical constituent structure. *Phonology*, 26(03):477–524.

Suprun, A. E. (1957). K upotrebleniju roditel'nogo i imenitel'nogo padežej množestvennogo čisla prilagatel'nyx v sočetanijax s čislitel'nymi *dva, tri, četyre* v sovremennom russkom jazyke. (On the use of the genitive and nominative plural adjectives in constructions with the numerals 'two', 'three', 'four' in Modern Russian.). *Učenye zapiski Kirgizskogo gosudarstvennogo pedagogičeskogo instituta (Proceedings of the Kyrgyz state pedagogical institute)*, 3:72–84.

Suprun, A. E. (1969). Slavjanskie čislitel'nye (stanovlenie čislitel'nyx kak osoboj časti reči) (Slavic numerals (establishing numerals as a separate part of speech)). Minsk, Belorussian

State University.

Szmrecsanyi, B. and Hinrichs, L. (2008). Probabilistic determinants of genitive variation in spoken and written English: A multivariate comparison across time, space, and genres. In Nevalainen, T., Taavitsainen, I., Pahta, P., and Korhonen, M., editors, *The dynamics of linguistic variation: Corpus evidence on English past and present*, pages 291 – 309. John Benjamins.

Tabak, W., Schreuder, R., and Baayen, R. H. (2010). Producing inflected verbs: A picture naming study. *The Mental Lexicon*, 5(1):22 – 46.

Tabak, W. M., Schreuder, R., and Baayen, R. H. (2005). Lexical statistics and lexical processing: Semantic density, information complexity, sex and irregularity in Dutch. In Kepser, S. and Reis, M., editors, *Linguistic Evidence — Empirical, Theoretical, and Computational Perspectives*, pages 529 – 555. Mouton De Gruyter, Berlin.

Thornton, R. and MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language*, 48:740–759.

Tily, H., Gahl, S., Arnon, I., Snider, N., Kothari, A., and Bresnan, J. (2009). Syntactic probabilities affect pronunciation variation in spontaneous speech. *Language and Cognition*, 1(2):147–165.

Tily, H. and Kuperman, V. (2012). Rational phonological lengthening in spoken Dutch. *Journal of the Acoustic Society of America*, 132(6):3935–3940.

Timberlake, A. (2004). *A Reference Grammar of Russian*. Cambridge University Press, Cambridge, U.K.

Tremblay, A. and Tucker, B. V. (2011). The effects of N-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon*, 6(2):302–324.

Turk, A. and Shattuck-Hufnagel, S. (2000). Word-boundary-related duration patterns in English. *Journal of Phonetics*, 28:397 – 440.

Ullman, M. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30(1):37 – 69.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., and Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, 9(2):266 – 276.

Ullman, M. T., Pancheva, R., Love, t., Yee, E., Swinney, D., and Hickok, G. (2005). Neural correlates of lexicon and grammar: Evidence from the production, reading and judgment of inflection in aphasia. *Brain and Language*, 93:185 – 238.

Van Lancker-Sidtis, D. and Rallon, G. (2004). Tracking the incidence of formuform eexpression in everyday speech: methods for classifications and verification. *Language & Communication*, 24:207 – 240.

van Son, R. and Pols, L. C. W. (2003). How efficient is speech? *Proceedings of the Institute of Phonetic Sciences*, 25:171–184.

van Son, R. and van Santen, J. (2005). Duration and spectral balance of intervocalic consonants: A case for efficient communication. *Speech Communication*, 47:100–123.

van Son, R. J. J. H., Bolotova, O., Lennes, M., and Pols, L. C. W. (2004). Frequency effects on vowel reduction in three typological different languages (Dutch, Finnish, Russian). In *INTERSPEECH-2004*, pages 1277–1280.

Vigliocco, G., Butterworth, B., and Garrett, M. F. (1996a). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, 61:261–298.

Vigliocco, G., Butterworth, B., and Semenza, C. (1995). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, 34:186–215.

Vigliocco, G. and Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin*, 128:442–472.

Vigliocco, G., Hartsuiker, R. J., Jarema, G., and Kolk, H. H. J. (1996b). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes*, 11(4):407–442.

Vigliocco, G. and Nicol, J. (1998). Separating hierarchical relations and word order in language production: is proximity concord syntactic or linear? *Cognition*, 68(1):B13 – B29.

Vitevitch, M. S. and Luce, P. A. (2004). A Web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*, 36:481–487.

Walsh, T. and Parker, F. (1983). The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics*, 11:201–206.

Weiner, E. J. and Labov, W. (1983). Constraints on the Agentless Passive. *Journal of Linguistics*, 19(1):pp. 29–58.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustic Society of America*, 91(3):1707 – 1717.

Zipf, G. K. (1935). *The Psycho-biology of language: An Introduction to Dynamic Philology*. Houghton Mifflin, Boston.

Appendix A

Stimuli and fillers for Chapter 2

A.1 Sg-Sg and Sg-Pl sentences

The pizza with the missing slice(s) looks unappetizing in the____
morning basket

The phone with the new keypad(s) works better now than when we first____
got it moved in

The truck with the special bumper(s) brakes unreliably in ____
heavy rain bad traffic

The ship with the wooden deck(s) leaks in high winds or rough____
seas weather

The desk with the sliding drawer(s) makes a lot of noise when ____
I'm working it's broken

The shark with the strong fin(s) likes playing and hunting around the____
aquarium reef

The plant with the delicious root(s) takes a lot of water____
every day to grow big

The stereo with the tiny switch(es) works best with a ____
good amplifier skilled technician

The loaf with the exotic grain(s) lacks all nutritional____
ingredients value

The telescope with the polished lens (es) picks up the faintest stars____
and galaxies very clearly

The fan with the wide blade(s) rocks on the highest____
shelf settings

The box with the dented corner(s) blocks everyone's way through____
the museum | the door

The statue with the imported stone(s) wrecks tourist pictures of____
the parade | the skyline

The beach with the sloping dune(s) bakes in the heat of____
the August sun | the circus tent

The rollerblade with the metal axle(s) locks easily onto____
one's sneakers | the bike rack

The movie with the famous scene(s) risks alienating its____
audience | investors

The satellite with the integrated computer(s) tracks all nearby____
radio waves | enemy satellites

The newsletter with the insightful article(s) backs unexpected____
candidates | policies

The report about the democratic senator(s) keeps appearing on all the ____
news channels | lawyers' desks

The picture of the jewel thief/thieves bumps all other news stories ____
off the air | into obscurity

The party for the fraternity brother(s) tops previous parties in ____
noisiness | neighbor complaints

The wedding of the congregation member(s) groups outside guests in____
a balcony | a separate pew

The secret about the movie star(s) keeps everyone guessing about ____
the truth | his lover

The song by the folk singer(s) leaps to mind whenever I think ____
about home | of the sixties

The verdict for the accused killer(s) locks up the accomplices____
for twenty years | to the crimes

The meeting for the college republican(s) wraps up a long series ____
of public lectures | of dinner parties

The protest for the political activist(s) banks on public support for____
the cause | lower taxes

The statement by the military terrorist(s) wrecks all hope of____
an end to conflict | a peaceful solution

The contract for the business executive(s) pipes tax money into ____ pensions investments

The agreement for the corporate lawyer(s) drops all reference to ____
liability mediation

The manuscript by the nuclear physicist(s) pumps excitement into ____
his colleagues the science press

The plan of the untraditional architect(s) maps out an eco-friendly vision for____
our town developers

The curse of the ancient mummy/mummies trips up unlucky ____
grave-robbers anthropologists

The article by the newspaper journalist(s) copes easily with ____
public objectsion counterarguments

The policy of the local politican(s) docks public salaries ____
by ten percent and raises taxes

The story about the religious leader(s) takes us from his childhood to his____
death marriage

A.2 Coll-Sg and Coll-Pl Sentences

The gang with the dangerous rival(s) camps out in interstate____
rest stops underpasses

The committee from the local union (s) tips bar tenders whenever____
they stay late the beer is good

The clan of the Scottish monarch(s) aches to reclaim the kingdom's ____
medieval glory sheep-breeding prize

The choir for the church service(s) picks out the songs without ____
rehearsing them the pastor's input

The audience at the tennis match(es) gasps at every unusually bad ____
serve outfit

The team in the shoe advertisement(s) winks at people in the____
grocery store sports arena

The class in the writing competition(s) types up all entries____
in Arial a day early

The staff for the training program(s) bakes eight types of pie for____
orientation the new students

The cast in the soap opera(s) laps up praise from critics____
eagerly shamelessly

The crew with the peacekeeping force(s) copes badly with the local____
customs resistance

The faculty with the research award(s) types up their results without any ____
assistance spell check

The fire brigade outside the building(s) pumps all the water from the____
fire hydrant duck pond

The student club in charge of the party/(ies) tacks up flyers all over____
campus town hall

The crowd at the Olympic event(s) lacks enthusiasm for the____
winner snack bar

The cleaning agency for the diabetes clinic(s) parks in the tow-away zone with ____
impunity great skill

The actors' guild in charge of the charity/(ies) tapes acceptance speechs on____
DVD awards night

The delegation from the wealthy foundation(s) taps previous donors with____
persistance many flyers

The clergy from the rural church(es) talks of aiding the____
poor and weak sermon writers

A.3 Sg-CollPl Sentences

The strength of the volunteer armies blocks other countries from____
attacking smuggling

The sight of the small villages kicks extra energy into the____
travelers brigands

The location of the pine forests marks a boundary between the____
continents campuses

The record of the soccer teams drops every time the new coach is____
present sick

The size of the protest groups leaps out at anyone watching the____
parade movie

The barricade erected by the crowds leaks ocean water despite the____ sand bags low tide

The strategy conceived by the battalions tricks enemy generals with little____
experience money

The territory defended by the packs wraps all around the town hall and____
courthouse market

The routine of the local militias risks boring the officers'____
wives men

The dance of the Navajo tribes tips autumn into winter when done____
badly just right

The boat waiting for the navy crews docks at many ports before ____
September refueling

The profit of the southern cartels tops previous records for smuggled____
contraband olive oil

The rhythm of the jazz bands rocks audience members into a____
stupor frenzy

The owner of the alpaca herds pipes clean water to the pastures from____
the pasture the river

The announcement made to the assemblies maps out a vision for the town's____
future parking lots

The future of the Russian brigades looks bleak after the multiple____
scandals blizzards

The destination of the Atlantic fleets tricks enemies into sailing too ____
far south quickly

The triumph of the northern navies bumps up public support for the new____
strategy admiral

A.4 D-CollPl Sentences

The crowd inside subway cars breaks up the commute by reading____
novels comic books

The jury in the folding chairs talks about the fight between the____
thugs judges

The crew in the drifting lifeboats makes a crude sail to catch____
fresh water jellyfish

The clan in the ancient castles books private tours for people____
who call ahead | have Scottish roots

The gang on the black motorcycles winks often when talking with____
the police | baristas

The hired help in the huge mansions tucks cashmere covers over the____
pillows | toilet seats

The squad inside the armored vehicles likes to read magazines during____
stake-outs | red lights

The team in the race cars marks a sponsor logo on all ____
car bumpers | merchandise

The staff in the tiny cubicles groups timecards according to____
last name | hair color

The family on the new bicycles parks anywhere when there are no____
bike racks | police cars

The party in the stretch limousines corks up only some of the____
bottles | kegs

The group on the jet skis blinks in surprise at the enormous____
shark fins | fishing boats

The troop inside the canvas tents kicks away large rocks before____
setting up camp | playing soccer

The orchestra on the tour buses camps under the stars after big____
concerts | parties

The audience in the luxury boxes gasps at the appalling ____
false note | costumes

The class at the lunch tables tapes up the rips in the____
science fair posters | white lab coats

The tribe inside stone huts cakes dust on top of food scraps because of____
flies | germs

The posse on the spotted horses drinks ale before setting off to find____
women | bad guys

A.5 Fillers

Many of the journalist's ____ don't know his sources' colleagues lovers

His father never ____ any of his childhood past-times
approves of takes part in

Children are constantly falling down, so don't ____ their skinned knees
overlook fuss about

Bark beetles ____ of deforestation in much of Canada
cause suffering from

Down comforters irritate the nasal passages of most ____
allergy sufferers poodles

A good cup of tea ____ most reasons for dissatisfaction in life
fixes hides

The lamp with the broken light bulb ____ as much light as the television
gives emits

Russian poets never ____ the praise that their work merited
receive acknowledge

Fleece slippers ____ much fuzzier on your feet than on the sheep's feet
feel look

When an alpaca ____ it looks like an alien invader
swallows spits

I have never ____ as much pride in the children as their first day of school
felt taken

One need only ____ the potion to understand its reputation
taste look at

The red cordury upholstery ____ your cheek when you sleep on it
imprints soothes

A silken blouse is a ____ house-warming gift for one's neighbor
sensuous sassy

It's hard to find a research assistant who ____ good records
keeps withholds

Tying one's shoes ____ hand-eye coordination in children
builds destroys

Fortified wheat flour cupcakes ____ surprisingly like cardboard
look feel

Spandex bike shorts cause much ____ among conference goers
laughter admiration

A motorcycle helmet improves ____ on the highway
safety style

A bottle of Elmer's glue has ____ nutritional value for the ingredients
suprising little

____ without a good outcome is not worth buying, let alone talking about
A mystery book An election

____ rarely interrupts our conversation unless dinner is especially late
Our son Our dog

____ has never bothered to look into how much a vacation would actually cost
The travel agent The department chair

____ doesn't believe in grading on a curve, despite the students' pleas
The professor The dance teacher

____ of the string quartet was enormously disappointing
The review the performance

After the ____ I have no appetite for the chicken dumpling dinner
cookies bad news

____ make an excellent breakfast, but an even better dinner

Bacon and eggs with sausage French toast with maple syrup

____ and heavy rain bring down power lines every November
The high winds The sunshine

____ without a hope of escape usually resigns himself to his fate
The prisoner The graduate student

____ with an ounce of brains can see that the politician is lying
Anyone An opossum

Are the seedlings planted near the corn growing too slowly?
Was the child's behavior towards her teachers appalling?
Do the springs and the cushions of the sofa creak every time you move?
Was the ugly large green tablecloth stained?
Did the bubbling, browning apple pie smell better than it tasted?
Did the free range turkey dinner satisfy anyone but Grandma?
Were the antiquated bodice and petticoats difficult to walk in?
Did the red and orange sunset impress the tourists?
Were the cars parked along the block towed away?
Did either the medicine or the surgery cure his kidney stones?
Did either the exam score or the homework assignments help the final grades?
Did either the high heels or the stockings go with her jacket?
Did both the makeup and the hairdo disguise her ugly dress?

did both the massage and the acupuncture help her back pain?
Did both the extra credit and the interesting lectures improve the professor's popularity?
Did both the computer and the smartphone lose the back-up data?
Did either the lunch or the dinner taste better than the breakfast?
Did either the wedding or the reception soften her father's attitude?
Were either the concert or the after-party more fun than a root-canal?
Are either the book or the movie better than the TV show?
Is a job without fulfillment worth having?
Is a book without a plot worth reading?
How is Brahms's fourth symphony when played on period instruments?
Can orchestras stay in tune without a good conductor?
Are the raccoons constantly raiding your trash cans?
Do elephants eat peanuts rarely or frequently?
Are circus clowns the scariest of all entertainers?
Do children misbehave without proper discipline?
Does Beethoven sound divine with the right violinist?
How did the cheese and scallion pie please your neighbors?

Appendix B

Stimuli and Fillers for Chapter 3

B.1 Critical sentences

Два журналиста __ (быть) в классе
Two journalists __ (be) in the classroom

Пять бульдозеров __ (появиться) на стройке
Five bulldozers __ (appear) on the construction site

Десять хулиганов __ (виднеться) в окне
Ten hooligans __ (be visible) through the window

Несколько памятников __ (оказываться) в центре внимания
Several monuments __ (turn out to be) in the center of attention

Мало хирургов __ (оставаться) в больнице
Few surgeons __ (remain) in the hospital

Много прудов__ (находиться) в парке
Many ponds __ (be located) in the park

Два рисунка __ (лежать) в коробке
Two drawings __ (lie) in the box

Пять преступников __ (покраснеть) от стыда
Five criminals __ (blush) from shame

Десять листьев __ (расти) на дереве
Ten leaves __ (grow) on the tree

Несколько акробатов __ (висеть) под куполом цирка
Several acrobats __ (hang) beneath the circus big top

Мало грузовиков __ (стоять) в проезде
Few trucks __ (stand) in the driveway

Много авторов __ (идти) по улице
Many authors __ (walk/go) along the street

Два партизана __ (пробиваться) к своим

Two partisans __ (get through to) their own people

Пять гектаров __ (давать) хороший урожай
Five hectares __ (give) a good harvest

Десять пленников __ (ударить) черную собаку
Ten prisoners __ (hit) the black dog

Несколько каналов __ (показывать) эту передачу
Several channels __ (show) this program

Мало спортсменов __ (показывать) отличные результаты
Few athletes __ (show) excellent performances

Много потоков __ (пробиваться) сквозь щели
Many rivulets __ (get through) the cracks

На полке __ (быть) два мобильника
On the shelf __ (be) two cell phones

На стройке __ (появиться) пять инженеров
On the construction site __ (appear) five engineers

За домом __ (виднеться) десять кустов
Behind the house __ (be visible) ten bushes

В центре внимания __ (оказываться) несколько писателей
In the center of attention __ (turn out to be) several writers

На улице __ (оставаться) мало автобусов
On the street __ (remain) few buses

У доски __ (находиться) много студентов
By the blackboard __ (be located) many students

На траве __ (лежать) два зоолога
On the grass __ (lie) two zoologists

В саду __ (покраснеть) пять помидоров
In the garden __ (turn red) five tomatoes

В одном доме __ (расти) десять мальчиков
In one house __ (grow) ten little boys

На вешалке __ (висеть) несколько пиджаков
On the coat rack __ (hang) several jackets

У постели __ (стоять) мало священников
By the bed __ (stand) few priests

По рельсам __ (идти) много поездов
Along the rails __ (walk/go) many trains

Красным цветом __ (писать) два самописца
Two automatic recording styluses __ (write) in red

Домашнюю работу __ (давать) пять преподавателей

Five teachers __ (give) homework

После оползня дом __ (ударить) десять камней
Ten rocks __ (hit) the house after the avalanche

По пятницам __ (работать) несколько поваров
Several cooks __ (work) on Fridays

Зимой хорошо __ (работать) мало лифтов
Few elevators __ (work) well in the winter

Статьи __ (писать) много профессоров
Many professors __ (write) articles

B.2 Sentences with end-stressed verbs

Два героя __ (спасти) деревню от дракона
Two heroes __ (save) the village from the dragon

Пять стульев __ (мочь) уместиться вдоль стола
Five chairs __ (can) fit along the table

Десять неудачников __ (найти) только разочарование в жизни
Ten unlucky people __ (find) only disappointment in life

Несколько пальцев __ (плести) небрежные косы
Several fingers __ (weave) untidy braids

Мало пастухов __ (вести) стадо от реки
Few shepherds __ (lead) the flock away from the river

Много тракторов __ (везти) груз по дороге
Many tractors __ (convey) the load along the road

При торнадо местныдва жрителей __ (спасти) два подвала
During the tornado two basements __ (save) the local residents

Заказы большой деревни __ (мочь) выполнять пять сапожников
Five shoemakers __ (can) fill the orders of a large village

В гавани __ (найти) убежище десять кораблей
Ten boats __ (find) shelter in the harbor

На ярмарке корзинки __ (плести) несколько ремесленников
Several artisans __ (weave) baskets at the fair

В тупик __ (вести) мало путей
Few paths __ (lead) to nowhere

Яблоки __ (везти) на рынок много ямщиков
Many coachmen __ (convey) apples to the market

B.3 Filler sentences

Больной отвечал на __ (вопросы) психологов
The sick man answered the psychiatrist's __ (questions)

большую часть домашнего __ (время) они проводили на ногах
They spent most of their domestic __ (time) on their knees

После гибели мужа барыня с __ (дети) поселились в Италии.
After the death of her husband the lady took up residence in Italy with her __ (children)

Около __ (деревья) ходила черная коров
A black cow was walking around the __ (trees)

Из шести __ (лаборантки) три были в отпуске
Out of six __ (female lab assistants) three were on leave

В отличие от __ (белка) бурундук не боится людей
As opposed to the __ (squirrel), the chipmunk is not afraid of people

Благодаря __ (компьютер) дети смогут легче работать дома
Thanks to the __ (computer) children can work at home more easily

С самой утренней __ (заря) дождь не переставал
The rain had not stopped since the __ (dawn)

На __ (крылья) своих жаворонки унесли капли росы.
On their __ (wings) the skylarks carried drops of dew

Девочка начинает искать __ (коробок) со спичками
The little girl began to search the __ (box) of matches

Легкий ветер время от __ (время) надувает занавески
From time to __ (time) a light wind blew the curtains

Поручик дал __ (лакей) целых пять рублей
The lieutenant gave the __ (footman) five whole rubles

Одной из самых известных историй о любви является история __ (барышня) и хулигана
One of the most famous love stories is the story of the __ (lady) and ruffian

Мальчик бежит вдоль __ (дорога)
The little boy runs along the __ (street)

Жена уговаривает __ (муж) навестить соседа
The wife persuades her __ (husband) to call upon the neighbor

Сюжет романа выстроен вокруг любовных историй двух __ (сестры)
The plot of the novel is built around the love story of two __ (sisters)

Братья проводят __ (зима) в доме тети
The brothers spend the __ (winter) in their aunt's house

Все письма остаются без __ (ответ)
All the letters remain unanswered [literally: without __ (answer)]

Литераторы пишут множество __ (письма)

The literary men write a great many __ (letters)

Изначально люди не обладали никакой __ (письменность)
Initially people did not have any kind of __ (written language)

Антонимы стали __ (предмет) лингвистического анализа сравнительно недавно
Antonyms became a subject of linguistic analysis comparatively recently

В античности город имел две __ (гавань), военную и торговую
In classical times the city had two __ (harbors), military and commercial

Лучше час свободы, чем сорок лет __ (тюрьма) и рабства
It is better to have an hour of freedom than forty years of __ (prison) and slavery

Свет __ (ракета) заполнил подвал
The light of the __ (rocket) filled the basement

Капитан уселся подле __ (офицер)
The captain sat down next to the __ (officer)

Собаки с __ (радость) побежали вперед
The dogs ran forward joyfully [literally: with __ (joy)]

Обычно к концу __ (лето) дачникам надоедало отдыхать
Usually towards the end of the __ (summer) the vacationers got sick of relaxing.

В __ (библиотека) не надо громко говорить
In the __ (library) you shouldn't talk loudly

В __ (коридоры) было почти так же жарко, как было холодно на улице
it was almost as hot in the __ (corridor) as it was cold outside

Над этой __ (проблема) действительно следует подумать
It's genuinely necessary to think about this __ (problem)

В буддизме нет __ (ненависть)
There is no __ (hate) in Buddhism

Эта область наиболее богата __ (астероиды)
This area is most rich in __ (asteroids)

Начало __ (мероприятие) пришлось перенести на два часа
It was necessary to move the beginning of the __ (activity) to two o'clock

Без __ (микробы) была бы невозможна жизнь на планете
Without __ (microbes) life on the planet would have been impossible

Всегда полезно знать, что делается в __ (стан) врагов
It is always useful to know what is being done in an enemy __ (state)

Опять надо было наведаться за __ (граница)
Again it was necessary to go on a visit abroad [literally behind __ (border)]

Нужно было срочно принимать __ (решение)
It was necessary to make a __ (decision) quickly

Напрасно было утешать __ (старуха)

It was useless to comfort the __ (old woman)

Он никогда не напишет __ (такой) письма
He will never write such __ (letters)

Командир гордится __ (свой) подчиненными
A commander is proud of __ (his) subordinates

Он был в синем костюме и __ (нейлоновый) рубашке
He was in a dark blue suit and __ (nylon) shirt

В теплом баре пахло __ (крепкий) кофе
In the warm bar it smelled of __ (strong) coffee

Отец принес кастрюлю со __ (свежий) икрой
The father brought a saucepan with __ (fresh) caviar

Из темноты бил в лицо __ (сильный) ветер
From the darkness a __ (strong) wind beat at one's face

Пароход подплывает к __ (небольшой) пристани
The steamship sailed up to a __ (smallish) dock

Извозчики остановились возле __ (освещенный) подъезда
The cabmn stopped next to an __ (illuminated) front door

На углу была __ (фотографический) витрина
On the streetcorner corner was a photographer's [literally: __ (photographic)] shop window

Одна шпилька лежала на __ (ночной) столике
One hairpin lay on the bedside table [literally: __ (nocturnal) table]

Агент по сбору объявлений присвоил три тысячи __ (казенный) денег
The advertising agent embezzled three thousand __ (government) [money units]

Ведомство оказалось в состоянии __ (системный) кризиса
The department turned out to be in a state of __ (systemic) crisis

Шея у жирафов необычайно __ (длинный)
The giraffe's neck is unusually __ (long)

В Германии существуют несколько __ (профессорский) должностей
In Germany there are several __ (professorial) duties

Военный фольклор богат __ (занимательный) рассказами
Martial folklore is rich in __ (entertaining) stories

Дядя начинает подозревать __ (тайный) помолвку
The uncle began to suspect a __ (secret) engagement

Старый друг выглядит __ (несчастный)
The old friend appears __ (unhappy)

Наконец правда о его __ (истинный) характере выплывет наружу
Finally the truth of his __ (true) character will come to light

Младшая дочь пытается привлечь внимание к __ (свой) персоне

The younger daughter tries to attract attention to __ (her) own self

Буквы международного __ (фонетический) алфавита подразделяются на три категории
The letters of the international __ (phonetic) alphaet are divided into three categories

В __ (западный) культуре псевдонимами пользуются только литераторы
In __ (western) culture pseudonyms are used only by literary people

Греческий – один из древнейших __ (письменный) языков мира
Greek is one of the ancient __ (written) language of the world

Чтение на __ (иностранный) языке труднее чтения на родном
Reading in a __ (foreign) languge is harder than reading in one's native language

__ (бедный) библиотекарша охала и качала головой
The __ (poor) librarian sighed and shook her head

Старик хотел сделать всех людей __ (счастливый)
The old man want to make all people __ (happy)

Талантливый драматург описал моменты из жизни __ (российский) монархов
The talented playwright described moments in the life of __ (Russian) monarchs

Для __ (нормальный) жизни достаточно зарабатывать 10 тысяч в месяц
Earning 10 thousand a month is enough for a __ (normal) life

__ (Цветочный) горшков на подоконнике не хватает
There are not enough __ (flower) pots on the windowsill

__ (Каждый) человеку хочется быть уверенным хоть в чем-то
__ (Each) person wants to be confident at least in something

На __ (свежий) воздухе всем спится лучше
Everyone sleeps better in __ (fresh) air

В __ (такой) условиях добиться успеха сложно
Achieving success in __ (such) conditions is complicated

В __ (любой) время года возможны ливни
Downpours are possible at __ (any) time of year

Самый обычный обед можно превратить в __ (праздничный)
It is possible to turn the most ordinary meal into a __ (holiday) meal

Требовалось точно сохранить все __ (цветовый) оттенки скульптуры
It was necessary to faithfully preserve all the sculpture's __ (colorful) nuances

Сейчас очень трудно найти __ (хороший) медсестер и фельдшеров
It is very hard now to find __ (good) nurses and medical assistants

Депутаты должны будут определить __ (свой) отношение к законам
The deputies will have to specify __ (their) attitude towards the laws

Сыщикам удалось обезвредить __ (жестокий) банду
The detectives managed to render the __ (vicious) gang harmless

Эту проблему нельзя решать по __ (один) алгоритму

One can't solve this problem with __ (one) algorithm

Девушка уверена, что любовник __ (презирать) ее кошку.
The young woman was certain that her lover __ (despise) her cat

Мама девочки __ (выйти) замуж по любви
The little girl's mother __ (get married) for love

Когда отец __ (умирать), его имение переходит к его сыну.
When the father __ (die) his estate passes to his son

Женщина __ (бояться), что князь больше ничего к ней не чувствует
The woman __ (be afraid) that the prince no longer felt anything for her

Цвет мёда __ (зависеть) от растений
The color of honey __ (depend) on the plant

Большое количество слов в европейских языках __ (иметь) латинское происхождение
A large quantity of words in European languages __ (have) a Latin origin

С момента создания этот закон __ (претерпеть) несколько переработок
From the moment of its creation this law __ (endure) several revisions

Люди __ (разговаривать) и одновременно посматривают телевизор
People __ (chat) and watch television at the same time

Мать и дочь __ (любить) друг друга безгранично
The mother and daughter __ (love) each other without limits

Наука о звуках речи __ (называться) фонетикой
The science of speech sounds __ (be called) phonetics

Взгляд Медузы __ (обращать) человека в камень
The gaze of the Medusa __ (turn) a man to stone

На дороге и на поле __ (светиться) месяц
The moon __ (shine) on the road and field

Фармацевтические фирмы __ (участвовать) в научных гонках
Pharmaceutical firms __ (participate) in scientific races

Католические монахи и монахини __ (покидать) монастыри
Catholic monks and nuns __ (abandon) the monasteries

Лошади очень хорошо __ (чувствовать) приближение грозы
Horses __ (feel) an approaching storm very well

Внуки и правнуки __ (беречь) землю русскую от врагов
The grandchildren and great-grandchildren __ (guard) Russian land from enemies

Собаки всю мебель __ (грызть)
Dogs __ (chew) all the furniture

Сани все время __ (скрести) полозьями и скрипят
All the while the sleighs __ (scrape) their runners and creaked

Враги не __ (хотеть) смотреть друг на друга

The enemies did not __ (want) to look at each other

Соседи громко кричали и__ (плакать)
The neighbors shouted and __ (cried) loudly

B.4 Intervening filler words

Adverbs	Gloss	Verbs	Gloss	Nouns	Gloss
аккуратно	carefully	арестовать	arrest	акула	shark
безопасно	safely	бежать	run	акушерка	midwife
внимательно	attentively	беседовать	chat	архитектура	architecture
достаточно	sufficiently	вглядываться	peer at	введение	introduction
естественно	naturally	восхищаться	admire	винт	screw
заботливо	thoughtfully	встречать	meet	выведение	removal
заметно	noticeably	выворачивать	unscrew	деревья	trees
заумно	overly abstrusely	высиживать	brood (as a hen)	детство	childhood
значительно	considerably	выскребать	rake out	истерика	hysterics
издалека	from far away	готовить	prepare	карандаши	pencils
инкогнито	incognito	ехать	drive	кузен	cousin
испуганно	fearfully	жалеть	pity	ландшафт	landscape
коротко	shortly	жаловаться	complain	лоджия	loggia
красиво	beautifully	зависить	depend	любовь	love
легко	easily	заворачивать	tighten	любопытство	curiosity
ловко	adroitly	задавать	give	местность	locality
медленно	slowly	исполнить	fulfill	молодость	youth
мрачно	gloomily	кататься	go for a ride	налог	tax
мучительно	agonizingly	лазать	climb	напиток	drink
мысленно	mentally	нажать	press	ненависть	hatred
надолго	for a long time	перебегать	run across	опера	opera
напрасно	in vain	подключать	connect	площадка	platform
невнятно	inarticulately	подчиняться	obey	прерогатива	prerogative
неразборчиво	unintelligibly	предпочитать	prefer	приязнь	goodwill
очаровательно	charmingly	рисовать	draw	спички	matches
практично	practically	сжигать	burn	статуя	statue
расторопно	deftly	сидеть	sit	таракан	cockroach
скучно	boringly	склониться	incline	тигр	tiger
стыдно	shamefully	скользить	glide	убежище	haven
хорошо	well	собирать	collect	условие	condition
чутко	keenly	спасаться	escape	фасад	façade
ясно	clearly	танцевать	dance	цветение	flowering
заранее	in advance	зарабатывать	earn	заседание	meeting
кстати	by the way	защищать	defend	зверь	beast
наверно	probably	кивнуть	nod	карандаш	pencil
ласково	tenderly	казаться	seem	идиот	idiot

Appendix C

Stimuli and fillers for Chapter 4

C.1 Critical sentences (list A)

На __ (экскурсия) в музей пошло мало студентов
Few students went on the trip to Moscow

На __ (собрания) всегда шли много споров
Many arguments always went on at the meetings

Вдоль __ (дорога) росли много кустов малины
Along the road grew many berry bushes

Наших __ (важный) гостей развлекло много актеров
Our important guests were entertained by many actors

В большом доме на __ (соседний) улице выросло много детей
In the big house on the neighboring street grew up many children

Мало сотрудников пошли на вечернее собрание __ (этот) учреждения
few coworkers went for an evening meeting of the company

Мало героев привлекли меня в романах __ (Достоевский)
Few heros attracted me to Dostoevsky's novels

Мало мальчиков росло без отца в этой __ (страна)
Few boys were growing up fatherless in in this country

Книги в __ (библиотека) прочло мало аспирантов
The books in the library were read by a few grad students

Много кранов текли на пятом этаже __ (высотный) здания
Many faucets were leaking on the fifth floor of the high-rise building

В этот __ (момент) его привлекло несколько городов
At this moment several cities attracted him

Первоклассников __ (этот) школы развлекли несколько концертов
The first graders of this school were entertained by several snowballs

В __ (пустыне) росли несколько кактусов

In the desert grew several cacti.

Около его __ (дача) текло несколько ручейков
Near his dacha flowed several little streams

В __ (лес) недалеко отсюда цвело несколько апельсинов
In the forest, not far from here, bloomed several orange trees

По стоянке __ (наш) дома шло несколько слепых
In the parking lot of our house walked several blind men

На __ (кровать) больницы легли несколько пациентов
Several patients lay down in the hospital beds

Много маяков привлекло туристов к западному __ (берег) Ирландии
Many lighthouses attracted tourists to the western coast of Ireland

Много товаров привлекли покупателей в __ (магазины).
Many goods attracted buyers into the magazines

Много видов развлекли пассажиров при __ (переезд) через горы
Many views entertained the passengers during the mountain crossing

Много потоков текло через сад в __ (река)
Many rivulets flowed across the garden into the river

За его __ (дом) текли пять труб
Behind his house leaked five pipes

К нашему __ (удивление) на забытой клумбе цвело пять ирисов
To our amazement, in the forgotten garden bloomed five irises

Много туристов брели по __ (улицы) Сан Франциско
Many tourists trudged the streets of San frantsistso

Много родителей шло на наши __ (собрания)
Many parents went to our meetings

Много учеников легли на влажной траве около__ (озеро)
many students lay on the wet grass by the lake

Много солдат ползло через дорогу к __ (убежища)
Many soldiers crawled across the road toward the refuge

Много детей уползли под __ (стол) за ужином
Many children crawled under the dinner table

После шторма __ (разбитый) стекло и мусор мело много волонтеров
After the storm many volunteers swept up the broken glass and rubbish

Грязный __ (пол) печально мели много рабочих
Many workers gloomily swept the dirty floor

Снаряжение на __ (поле) несли много спортсменов
Many athletes brought equipment to the field

__ (Гречишный) блины и лепешки пекли много поваров в нашем городе

Many cooks in our town baked buckwheat blini and cakes

__ (Бездомный) людям охотно помогло много добровольцев
Many volunteers willingly helped the homeless people

Эти __ (мудрый) слова произнесло много стариков
Many old men pronounced these wise words

В __ (голодный) годы мало фермеров везли большой урожай на рынок
In hungry years few farmors brought their harvest to market

При __ (такой) грозе через реку гребли мало рыбаков
During such a storm few peasants would row across the river

Мало крестьян плели новые лапти на __ (лето)
Many peasants wove new bast sandals for the summer

Мало аспирантов пренебрегло профессорами того __ (университет)
A few graduate studentes disdained the professor of this university

Мало охранников стерегло диких зверей в __ (зоопарк)
A few guards guarded the wild animals at the zoo

Через __ (узкий) долину медленно текло два ледника
Across the valley flowly slowed two glaciers

Вдоль __ (канал) уныло брело пять бедняков.
Along the canal five paupers trudged gloomily

По __ (улица) Берлина шли пять солдат
Along the streets of Berlin walked five soldiers

К __ (мы) домой быстро шли десять известных политиков
Ten important politicians walked quickly to our home

Несколько папоротников росло напротив детского __ (садик)
Several ferns grew across from the kindergarten

После экзамена несколько студентов легло на __ (трава)
After the exam several students lay down on the grass

На __ (пляж) несколько младенцев ползли по песку.
On the beach several small boys crawled along the sand

Мебель к __ (мой) дому привезло пять грузовиков
Five trucks brought the furniture to our home

Много отрядов вело бои на окраине __ (город)
Many troops brought war to the border of our city

Десять грибов росли на поляне, где __ (играть) дети
Ten mushrooms grew on the hill where children were playing

Десять ясеней росло на высоком __ (холм)
Ten apple trees grew on the high hill

Десять тюльпанов цвели между нашим и __ (соседний) домом

Ten tulips bloomed between our house and the neighboring one

Той __ (ночь) костры жгли несколько туристов
That night tourists burned several bonfires

__ (Пыльный) коридоры мели несколько уборщиков
Several janitors swept the dusty corridors

__ (Животные) в наш город привезло несколько циркачей
Several circus men brought the animals to our city

Пять выводов влекло Бориса к __ (профессия) архитектора
Five conclusions attracted Boris to the profession of architecture

Пять бутонов цвели на розовом __ (куст)
Five buds bloomed on the rose bush

Нас к его __ (дом) привезло десять лимузинов
Ten limousines brought us to his house

__ (Окна) скребли десять кустов сирени.
Ten lilac bushes scraped at the window panes

Много школьников произнесли немецкие __ (слов) правильно
Many schoolboys pronounces the German words properly

Много крестьян толкло муку из сухих __ (ягода) толокнянки
Many peasants ground flour from dry manzanita berries

__ (Весь)тяжесть машины несло два домкрата.
Two jacks bore the entire weight of the car

Пять сыновей росли крепкими и __ (здоровый)
Five sons grew up strong and healthy

Десять юристов шло на встречу с __ (бастующий)
Ten lawyers went to meet the workers on strike

__ (Стол) в его столовой трясли десять хулиганов
Ten hooligans shook the table in his dining room

__ (Скучающий) гостью развлекло два приятеля
Two friends entertained the bored guest

Пять поездов везло пшеницу в __ (Москва)
Five trains carried grain to Moscow

Пять голосов произнесло клятву __ (врач)
Five voices pronounced the doctor’s oath

Несколько сторожей стерегли пленников в __ (тюрьма)
Several guards guarded the prisoners in prison

Десять автобусов привезли наших студентов к __ (музей)
Ten buses brought our students to the museum

Пять вождей пренебрегли предсказаниями __ (старейшина)

Five leaders neglected the elders' warnings

Во __ (двор) листья гребло два подростка.
In the courtyard two teenagers raked the leaves

Десять пастухов пасли белых овец на __ (граница) леса
Ten shepherds tended white sheep on the edge of the forest

Два инспектора шло через поле к __ (станция)
Two inspectors walked across the field to the station

Два матроса волокли дырявый ялик по __ (песок)
Two sailors dragged the holey boat along the sand

В киоске два продавца плело красно-синие __ (фенечка)
In the kiosk two sellers wove red and blue bracelets

Два пожарных спасли канарейку и __ (котенок) от пожара
Two firemen saved a canary and a cat from the fire

C.2 Filler sentences

Во второй половине дня мы __ (быть) свободны
We were free in the second half of the day

Мой отец, будучи __ (больной), вообще никуда не ездил
My father, an invalid, generally didn't go anywhere

В лесу __ (лежать) глубокий снег
In the forest lay deep snow

Мой старший брат похож на нашего __ (дядя)
My older brother looks like our uncle

Обстоятельства преступления были явно __ (отягчающий)
The circumstances of the crime were obviously aggravating

Мой отец __ (стать) известным театральным критиком
My father became a famous theater critic

Возле __ (плотина) был построен деревянный дом
Alongside the dam there was built a wooden house

За чаем __ (наш) бабушка была почти веселая
At tea our grandmother was almost cheerful

В последнюю __ (неделя) снегу подсыпало
Over the past week some snow has sprinkled down

Этих __ (пустяк) мне хватило на всю жизнь
I've had enough of such trifles for a lifetime

__ (Овощи) должно хватить на всю зиму

The vegetables are supposed to be enough for the whole winter

Никто не был безразличным к его __ (судьба)
No one was indifferent to fate

В таком городишке ничто не __ (остаться) секретом
In such a town nothing remains a secret

Вдоль тротуаров стояли __ (запаркованный) машины
Along the sidewalks stood parked cars

Из __ (палата) полководца звуков не доносилось
From the commander's tent no sounds carried

От нашей хибары не __ (уцелеть) и фундамента
Of our hut not even the foundation survived

Почти ничего с детства не __ (запомниться)
Almost nothing from childhood stayed in memory

Брата в три __ (час) не было дома
Brother wasn't home at three o'clock

Отец ежедневно __ (уходить) искать квартиру
Father went out every day to look for an apartment

Карта района лежала перед __ (мы) на столе
A map of the region lay in front of us on the table.

Спокойнее было на берегу Черного __ (море)
It was more peaceful on the shore of the Black Sea

Из соседнего дома __ (выйти) старуха
From the neighboring house came out an old woman

Власти преследовали тех, кто __ (посещать) церковь
The authorities persecuted those who went to church

Судьбу обвиненного __ (решать) его же сослуживцы
The fate of an accused person was decided by his own fellow workers

Он еще не __ (успеть) ничего обдумать
He still had not had time to think over anything

Эта работа интереснее, чем __ (какой)-либо другая
This work is more interesting than any other

Мать положила перед __ (каждый) по куску хлеба
Mother set on piece of bread in front of every person

Пароход __ (должно) прийти через два часа
The steamship is due to arrive in two hours

Нам было невозможно ехать на __ (трамвай)
It was impossible for us to go on the tram

Куприн дремал перед пустой __ (бутылка)

Kuprin dozed in front of an empty bottle

Ася была __ (уверен), что моя мама не согласится
Asia was convinced that my mother would not agree

Эмиграция может убить любого __ (писатель) в три года
Emigration can destroy any writer in three years

Грибы __ (любить) расти под этим деревом
Mushrooms love to grow under this tree

Учительница поставила __ (лампа) на стол
The teacher put a lamp on the table

Бедный сирота мечтал служить во французском __ (легион)
The poor orphan dreamed of serving in the French foreign legion

Полк __ (сосредоточиться) в лесу
The regiment was concentrated in the forest

Челнок мог доставить оборудование на __ (орбитальный) станцию
The shuttle was able to deliver equipment to the space station

Образцы __ (мочь) подготовить ассистенты
The assistants were able to prepare the specimen

Еноты могли навести __ (полный) хаос на кухне.
The raccooms wrought complete chaos in the kitchen

Выпускники от __ (радость) подбросили шляпы вверх
The graduates threw their hats up from joy

Мой бутерброд с __ (колбаса) съела мышь
A mouse ate my chicken sandwich

Молодой бизнесмен искал __ (комната) в гостинице
A young man was looking for a room in the hotel

Мой велосипед __ (купить) мне мой парень
My boyfriend bought me my bike

Маленькая девочка купалась в __ (большой) бассейне
A little girl was swimming in the big pool

__ (Красивый) девочка говорила по телефону с подругой
A pretty girl was talking on the phone with her friend

Ольга выпила слишком __ (крепкий) кофе
Michelle drank coffee that was too strong

__ (Мои) бижутерию украли хитрые воры
Tricky robbers stole my jewlery

Обезьяна __ (играть) в шахматы с дрессировщиком.
The monkey played chess with the trainer

На ее шее __ (висеть) серебряный амулет

On her neck hunt a silver amulet

Удача __ (зависеть) только от него самого
Success dependent only on him, himself

Сегодняшний обед __ (приготовить) их мать
Today's dinner was cooked by their mother

Мария купила очень дорогую __ (одежда)
Maria bought very expensive clothes

Она надела серую шляпу в __ (цветочек)
She wore a grey hat with little flowers

Ваши __ (большой) часы испортил Борис
Boris broke your big clock

Он получил __ (письмо) из библиотеки
He received a letter from the library

Его тетя заставила __ (он) постричься
His aunt forced him to get a haircut

Медсестра ухаживала за больными __ (дети)
The nurse took care of the sick children

Писатель подписывал __ (свой) книги
The writer was authographing his books

Эту __ (поэма) написал очень талантливый человек
This poem was written by a very talnted writer

Милиционеры __ (арестовать) протестующих
The cops arrested the protestors

В корзине лежали __ (красный) яблоки
In the basket were laying red apples

Лариса познакомилась с __ (новый) людьми
Larisa met new people

На заборе весь __ (день) сидела кошка
A cat was sitting on the fence all day long

Зайцы бежали в __ (сторону) леса
The spiders were running towards the forest

Оленя убил __ (ловкий) охотник
The deer was killed by a clever/quick hunter

Бабушка варила кашу для __ (внук)
The grandmother cooked meals for her grandson

Банкиры __ (спрятать) деньги в секретном сейфе
The bankers hid the money in a special safe

Мальчики __ (смотреть) в зеркало очень долго

The boys looked in the mirror for a while

На __ (стена) висел значок, обозначающий выход
On the wall was hung a sign, indicating the exit

Ученик поинтересовался, как достичь __ (бессмертие)
The student inquired about immortality

Дедушка выжал апельсиновый __ (сок)
The grandfather squeezed orange juice

Собака __ (смотреть) на него издалека
The dog was looking at him from far away

На стол залез __ (рыжий) кот
A red cat climbed on the table

Опытный летчик удачно __ (приземлить) самолет
An experienced pilot landed the plane successfully

Александр Грэм Белл __ (изобрести) телефон
The phone was invented by by Alexander Graham Bell

Пушкин __ (написать) поэму "Руслан и Людмила"
Pushkin wrote the poem "Ruslan and Ludmila"

Малину в лесу __ (собирать) милые дети
Children picked the raspberry in the forest

Кошелек __ (выпасть) из ее большой сумки
The wallet fell from her big purse

Кухню после рождественского ужина __ (убирать) все вместе
Everyone together cleaned the kichen after the Chirstmas dinner

Женщина без конца __ (махать) руками
The women constantly waved her arms

Его стихи читал __ (пожилой) человек
An elderly person read his poems

Портфель стоял в __ (угол) комнаты
The backpack stood in the corner of the room

Обеденный стол поместился в __ (столовая)
The lunch table fit into the dinning room

Квартиру купила молодая __ (пара)
The room was bought by a young couple

В его комнату свет не __ (попадать)
Light didn't get into his room

Надежда появилась в __ (самый) трудную минуту
Hope appeared in the most difficult moment

На концерт прилетели поклонники со __ (весь) мира

Fans from all over the world came to the concert

Цыганка __ (предсказать) его будущее
The gypsy predicted his future

__ (Грязный) белье стирала ее дочка
Her daughter washed the dirty laundary

Ее одежда соответствовала __ (погода)
Her clothes were appropriate for the weather

Мячик на дороге __ (выронить) дети
The ball on the road was dropped by the children

Судья __ (объявить) смертный приговор
The judge announced the death sentence

Воспаленный зуб __ (удалить) зубной врач
The dentist took out the infected tooth

Мать переживала за своих __ (дети)
The mother was worried about her children

Книги __ (уничтожать) революционеры
The revolutionists destroyed the books

Катя с трудом добралась с __ (вокзал) домой
Katya, with great difficulty, got home from the train

Марка вчера __ (обокрасть) бездомный
A homeless person robbed Mark yesterday

Он прожил всю жизнь на __ (Дальний) Востоке
He lived his whole life in the Far East

Мой племянник прилетел вчера __ (утро) из Северной Африки
My nephew flew in yesterday morning from North Africa

Моя бабушка великолепно танцевала __ (народный) танцы
My grandmother dances folk dances wonderfully

Жена Андрея сильным голосом __ (петь) прекрасные песни
Andrey's wife sang wonderful songs in a strong voice

Ночью разобрали железнодорожный __ (путь)
At night they dismantled the railroad

Враг отступал беспорядочными __ (кучка)
The enemy retreated in disordered clumps

Художник рисовал картины __ (выдуманный) городов
The artist drew pictures of imaginary cities

Летним утром приятно гулять по __ (роса)
On summery mornings it's pleasant to stroll when the dew is still on the ground

Все детали в __ (этот) цехе изготовляли роботы

All the details in this workshop were manufactured by robots

Отсюда __ (быть) хорошо виден нескошенный луг
From here the unmowed meadow was easily visible

На картине была изображена женщина с __ (длинный) волосами
In the picture a woman with long hair was portrayed

На первом этаже __ (помещаться) фотография
On the first floor was a photography studio

У __ (птица) оказалось сломанным крыло
The bird turned out to have a broken wing

На столе стоял букет красных и__ (белый) роз
On the table stood a bouquet of red and white roses

Капитан решил спустить с корабля __ (шлюпка)
The captain decided to lower the lifeboat from the ship

К __ (бутылка) тонкой проволочкой была прикручена пробка
A cork was tied to the bottle by a thin wire

Нас встретила девочка с __ (бантик)
We were met by a little girl with a bow

Она испугалась __ (слово) врача
She was frightened by the doctor's words

На телеге __ (стоять) бочонок с водой
On the cart stood a little keg of water

На его лице __ (выразиться) удовольствие
On his face was an expression of pleasure

Изредко мне случалось посещать __ (родной) места
I have rarely happened to visit the places of my birth

На форум __ (прибыть) руководители комитетов
The committee leaders were present at the forum

Таня опустила __ (сумка) на пол и заплакала
Tanya let her bag drop to the floor and began to cry.