The goal of this dissertation is to advance the state of the art of research in constraint-based phonotactics. It takes a two-pronged approach: a technological contribution intended to facilitate future research, and experiments which seek to shed light on high-level questions about the properties of phonotactic models that can guide the development of theoretical work.

The technological contribution is a software package called Speriment which allows experimenters to create and run experiments over the internet without advanced programming techniques. This software is particularly well suited to the kinds of experiments often run in phonotactic research, but can also be used for experiments in other domains of linguistics and the social sciences. It is hoped that this software will make it faster and easier to conduct phonotactic and other experiments as well as encourage experimenters to increase the reproducibility and transparency of their research.

The experiments presented here address questions are theory- but not framework-neutral. That is, they apply to constraint-based frameworks for theories of phonotactics, with the first study seeking to distinguish between two such frameworks, Harmonic Grammar and Maximum Entropy, while the second investigates whether phonotactic knowledge is independent of knowledge of phonological alternations. These coarse-grained questions about phonotactic knowledge — how pieces of phonotactic knowledge interact with each other and with another part of the grammar — are intended to add to the groundwork on which phonotactic models and models of all phonological knowledge are built. Their findings have implications for which constraint-based frameworks should be used to couch future theories and how these theories can be reliably tested.

The idea that experimentation can improve our understanding of phonotactic knowledge is central to this dissertation. Although the value of experimentation is largely recognized within the field of phonology, it is informative to look at the ways in which it advances our understanding of the questions addressed here. Despite calling into question the need for syntacticians to conduct experiments rather than informally collecting intuitions, **?**, points out that experiments are useful in cases where questions are too subtle to be satisfactorily answered by introspection. Experiments 1 and 2 in Chapter **??** are useful in demonstrating exactly how such cases can arise.

An intuition in phonology takes the form of a yes or no answer to whether a word is acceptable in a given language. I will call these first-degree intuitions. Graded first-degree intuitions are also possible, describing the degree to which a word is acceptable. Given two graded first-degree intuitions, we can ask which word is the more acceptable of the two. We can call this a second-degree intuition. In order to address the question in Chapter **??** of how constraint violations combine, we need third-degree intuitions: a comparison between second-degree intuitions which are themselves a comparison between first-degree intuitions. Linguists disagree about the consistency of first-degree intuitions in syntax (**?**). Second- and third-degree intuitions are even more difficult to access. So in addition to the other virtues of experiments described by Gibson and Fedorenko, we should recognize that experimental designs and quantitative methods allow us to use first-degree intuitions to build answers to second- or third-degree questions. Specifically, in Experiments 1 and 2, instead of asking a

few people whether [tlavb] is worse than [plag] to the same degree as the sum of the differences between, on one hand, [plag] and [tlag], and, on the other hand, [plag] and [plavb], we ask many people for primary intuitions on many items and quantitatively assess whether their answers collectively suggest that the differences are equal. The difficulty in second- and third-degree intuitions lies in their comparative nature. Quantitative methods allow us to do the comparison externally, across participants and items instead of within one person's mind.

First-degree intuitions are most directly useful in gathering evidence for constraints or rules — the content of a theory of phonotactics. First-degree intuitions that [tlag], [tlib], and [tlun] are unacceptable lend support to the theory that the constraint set in a constraint-based grammar contains a constraint like *tl, but, on their own, say nothing about the relationship of the constraint set to other parts of the grammar. This dissertation uses experimental designs to question the relationship of constraints to Eval and the relationship of constraints used in alternations to those used in phonotactics. It is perhaps not surprising, then, that the theory-neutral questions addressed here rely on fairly complex experimental designs.

Chapter **??** requires experimentation for a more obvious reason: it applies different treatments to different groups in order to look for a causal relationship between treatment and subsequent behavior. While again eliciting only first-degree intuitions, those intuitions stand to be affected by the training period. Our only hope of being confident that the effect is causal is to employ controls, by using two treatments and looking for a difference between their outcomes. This is of course a common use for experiments, central to artificial language learning studies as well as others.

For these reasons, experiments are well-suited to the questions addressed here as well as many others in phonological theory, and tools to facilitate their use are likely to remain in high demand. This dissertation explores ways to more efficiently create and share experiments and their results, as well as ways of using experiments to probe the structure of phonotactic knowledge.

In Chapter 1, I discuss Speriment's features and how it is used as well as compare it with other similar packages, each of which have different strengths and weaknesses. In Chapter 2, I review literature that suggests that any accurate model of phonotactics must allow for the accumulation of violations, so that the grammaticality of a word depends on all of the violations it contains. I present an experiment with the ability to refine this statement by investigating how additional violations affect the grammaticality of the word, weighing on the question of whether linear Harmonic Grammar or Maximum Entropy grammar is a better model for phonotactics. In Chapter 3, I consider whether a constraint may have a higher weight in the phonotactic grammar if it is active in alternations than if it is not. The effect of alternations is not usually accounted for in phonotactic models but may be quite important.