Scientific Standards for Linguistics

Daiki Hashimoto
University of Tokyo

1 Introduction

It is sometimes confusing for master students (and probably even PhD candidates) to understand how to write a thesis in linguistics. This is because the logics employed in previous literature varies depending on a linguist, and occasionally they are not scientific at all. The aim of this paper is to clarify what scientific arguments should be (Section 2), and how the scientific arguments have been employed in linguistics (Section 3). Section 4 offers some tips to write a linguistic paper.

2 Scientific argument

2.1 *What is science?* As you know, linguistics is a field of science. Odden (n.d.) notes that "The purpose of science, broadly constructed, is to gain knowledge of the nature of reality, and rendering it in an objective form so that it can be understood and evaluated by others." This purpose is achieved by the construction of a good theory. A theory is defined as a systematic set of hypotheses that accurately describe both actual observations and potential observations in the universe. Actual observations are those that were already made, and potential observations are those that may be made in the future. That is, a theory must capture every single phenomenon, regardless of whether it was already observed or not. A hypothesis is usually defined as a simple statement that provides an explanation for why a particular phenomenon takes place. It is further discussed in the following sections how hypotheses can be formed and developed. Multiple key hypotheses contained in a theory interact each other to account for the reality, as shown below. That is, a theory is like a book of statements that explain the reason why a phenomenon in the universe occurs. A theory is supposed to model something in reality to account for phenomena in the universe. For example, a biological theory models an ecosystem to account for why the nature is so orderly, and a linguistic theory models a mental module to explain why language is so orderly.

A variety of theories have been posited in science, even though "the eventual goal of science is to provide a single theory that describes the whole universe." (Hawkins, 1988: 9) As discussed below, it is extremely difficult to construct a single theory accounting for the whole universe, and thus science is divided into many branches such as physics, biology, and linguistics. As scientific branches construct their own theories, the number of theories is a function of the number of scientific areas.

What is more complicated, even a scientific branch is crowded with two or more competing theories. For example, phonologists have postulated a rule-based theory (Chomsky & Halle, 1968), a constraint-based theory (Prince & Smolensky, 1993/2004), and an exemplar-based theory (Pierrehumbert, 2002). These competing theories all aim to model how human beings memorize pronunciations and process the mental representations to form production targets, and they account for why a linguistic phenomenon takes place. This is why Odden (n.d.) outlines one inviolable requirement and three desirable attributes of a theory to single out a good theory.

The inviolable requirement is Falsifiability (Popper, 1959). Odden (n.d.) notes that "statements of a theory must have specific knowable meanings," and originally Popper (1959: 18) notes that "it must be possible for an empirical scientific system to be refuted by experience." All the results obtained from research need to either verify or disprove a theory. It is not allowed for a theory to have no potential for falsification. As noted in Section 2.3, a hypothesis is too abstract to test directly, and thus scientists always test the prediction in order to argue for or against the hypothesis. If a theory cannot deduce a specific prediction in relation to a variable of interest, then it is no longer a theory. For example, if a meteorological theory deduces a noncommittal prediction like "it will be sunny or not today," it fails to lead to a specific prediction and is no longer a scientific theory.

In addition to this inviolable requirement, Odden (n.d.) explicitly states three desirable attributes of a theory:

* This short article highly benefits from the discussion with Jen Hay (Canterbury) and Beth Hume (Ohio) during my PhD. Their comments on my research are always insightful, and I always appreciate their professional guidance. It goes without saying that my thanks also go to many of my colleagues in the University of Canterbury and the University of Tokyo. I am grateful to Luke Cairns for his kindness to comment on the early version of this paper. This paper also benefits from the discussion with first-year students of the University of Tokyo, when I teach a class called English for Arts and Science (EfAS).

Proceedings of *AJL 4*

(1) Three desirable attributes of a theory
Correctness: the observations entailed by the theory correspond to reality.
Comprehensiveness: the theory covers all essential attributes that define the set of entities taken to be in the domain of the theory, and that domain is the greatest possible.
Aesthetics: the statements of the theory are simple.

As discussed in Section 3, a linguistic theory usually puts forward a particular prediction such as "a speaker may produce variant X when they are exposed to Y" and "a listener may perceive variant X given environment Y." If an observation that is made in research is in line with the prediction deduced by a theory, the theory satisfies Correctness. On the other hand, if the result of an experiment disagrees with the prediction, the theory violates this desirable attribute.

Comprehensiveness favours a theory that can capture a wider range of phenomena. This is an extreme example, but a theory is highly valued if it can capture not only linguistic phenomena but also physic phenomena and biological phenomena. Hawkins (1988: 11) notes that "It turns out to be very difficult to devise a theory to describe the universe all in one go. Instead, we break the problem up into bits and invent a number of partial theories. Each of these partial theories describes and predicts a certain limited class of observations." This is why, linguists only focus on how linguistic knowledge is formed and used in speech, and physicists only study relativity and quantum gravity. (I believe that the universe is actually very simple, and a variety of scientific fields might be unified in the future, the result of which is that a single theory might be established.)

Aesthetics is another desirable attribute of a theory. This may be well-known as "Occam's razor." That is, a theory should not be complicated without necessity, and it is ideal to construct a theory with a smaller number of arbitrary assumptions. A theory with a small set of simple hypotheses is aesthetically better in comparison with that with a larger set of complicated hypotheses.

The goal of science is to develop a single theory that satisfies these requirement and attributes, which is considered to reflect the reality of the universe. Needless to say, this goal has not been achieved yet, and there still remains a large amount of knowledge uncovered. (This is why the modern society still has a profession called scientist.)

Before concluding this section, I would like to note that every single theory is provisional, and it is just a set of hypotheses, which we can never prove. Even if a theory is in line with the results of millions of experiments, what we can say is that the theory is supported by the results for the moment. There is always possibility that an observation may be made in the future which may refute a theory. Given the counter-example, the theory has to be modified with additional hypotheses or even discarded. This will be discussed further in Section 2.3.

This section noted that the goal of science is to construct a good theory that accounts for the reason why a particular phenomenon takes place in the world. The rest of Section 2 illustrates what kind of logics is employed in science and how a theory is developed.

2.2 *Induction: descriptive generalization and hypothesis* Every theory is kicked off by an inductive logic. Inductive reasoning expands our knowledge based on observations. Oxford University Press (2007) defines "induction" as "The process of inferring or verifying a general law or principle from the observation of particular instances." That is, this logical process makes up an abstract statement on the basis of concrete examples. Although there are several types of induction, the current paper only discusses two of them: descriptive generalization and abduction.

Descriptive generalization is to find a pattern of a variable of interest given a sample from larger population (see McCarthy, 2008: Ch.2.1.2). Based on a small number of concrete examples, we offer a general statement with regards to the whole population. Imagine that you are interested in the colour of apples. If you observe 100 apples, you might reach the conclusion that "every apple is red." This is a descriptive generalization. Although you have not checked the whole population of apples (since it is literally impossible), you form a descriptive generalization in relation to apples. Similarly, linguists have posited a variety of descriptive generalizations in relation to phonological phenomena. For example, Wells (1982) generalizes that a coronal plosive phoneme /t/ in an unstressed syllable is phonetically realized as a flap [ɾ] in relation to English phonology.

Based on a descriptive generalization, scientists develop a hypothesis. As noted in the preceding section, a hypothesis is a statement that accounts for the reason why a phenomenon, which can be descriptively generalized, takes place. Forming a hypothesis is called abduction or abductive inference. In relation to the generalization of apple colour, we may form hypotheses that "trees aim to spread their seeds by inducing animals to eat their fruits" and "animals are likely to pay attention to red stuff in nature" (Washington State University, 2016). These hypotheses account for why apples are red in nature. With regards to flapping, we can form hypotheses that "speakers pay small effort to the pronunciation of unstressed syllables" and "flaps [ɾ] are pronounced with smaller amount of articulation effort as compared with fully stopped plosives [t]." These hypotheses explain why a coronal plosive phoneme is realized as a flapped variant in an unstressed syllable. The degree of abstraction in a theory can be increased using a set of hypothetical concepts systematically, with the result that the theory can capture a wider range of phenomena. In linguistics, this has been done using some concepts such as feature geometry (Clements & Hume, 1995), markedness constraints (Prince & Smolensky, 1993/2004), and exemplars (Pierrehumbert, 2002).

Finally, we would like to note that there is always a more or less logical leap when a statement is formed by inductive reasoning (Verkerk et al., 2015: 224). Let us consider the descriptive generalization of apple colour. We descriptively generalized that apples are red based on only 100 apples, but how about the colour of the 101st apple that may be observed in future study? It might be white or purple. Despite the fact that we have not explored the whole set of apples in the world, we descriptively generalized that apples are red. This is obviously a logical leap, that is, a gap in the argument. Generalizations and hypotheses formed through induction are not necessarily true. The renowned philosopher David Hume even claims that "induction cannot be rationally justified" (Okasha 2016: 22). This is why the statement that was posited through inductive reasoning has to be tested on and on. This unceasing testing is performed using another type of logics called deduction.

2.3 *Deduction, prediction, and revision* Scientific deduction is an educated guess based on a theory. A scientist puts forward a specific prediction using a theory. That is, deductive reasoning is an inference process from a set of abstract statements (i.e., hypotheses) to a concrete statement (i.e., prediction), while inductive reasoning is an inference process from a concrete example (i.e., observation) to a set of abstract statements.

As noted in the preceding section, the hypotheses formed by inductive reasoning are not necessarily true. This is why, scientists usually use the existing hypotheses of a theory to test the validity via deductive reasoning. One might wonder why we do not test a theory and the consisting hypotheses directly. The reason is very simple: they are too abstract to test directly. The hypotheses such as "language faculty has a markedness constraint called NoCoda" and "successfully perceived tokens are stored as exemplars" cannot be observed using either a microscope or a MRI scanner. What we can do is to deduce an educated guess based on the hypotheses, and test the concrete prediction. If a theory satisfies the inviolable requirement called Falsifiability, it always allows us to single out a specific prediction: "Given that hypothesis X is true, then we should observe phenomenon Y."

A hypothesis is sometimes a statement like "If X, then Y." For example, Darwin's natural selection hypothesis can be stated like "If a given animal is adapted to a given environment, then she can survive." This hypothesis accounts for the reason why the nature is so orderly. Scientists can test the hypothesis by deducing a specific prediction. Remember that we always need additional hypotheses or assumptions in order to deduce a prediction. (These additional statements may be called warrants in the terminology of Toulmin (1964: Ch.3)). For example, we can deduce a prediction that "aquatic animals are all good at swimming" using the natural selection hypothesis. This prediction is based on not only the natural selection hypothesis but also other assumptions such as "swimming is a useful skill under water" and "a useful skill is an adaptive advantage." In this way, theoretical prediction is always performed based on two or more statements (i.e., hypotheses and assumptions), which are components of a theory. A couple of good deductive examples in linguistics will be reviewed in Chapter 2.

After a specific prediction is deduced, the prediction is tested by experimental or non-experimental observations. In the case of the prediction about swimming ability of aquatic animals, scientists may rely on non-experimental observations, that is, go to rivers and the sea to check a certain number of types of aquatic animals.[1] After the investigation of the collected data, the scientists may reach a particular conclusion with regards to aquatic animals. If the conclusion agrees with the prediction, then the theory is supported. In this case, the theory should stay intact, and it may be worth testing in future study again. On the other hand, if the conclusion disagrees with the prediction, then the theory has to be revised with another hypothesis or even discarded. This is because the theory does not satisfy the desirable attribute called Correctness. It needs to be revised so that it can capture the unpredicted result. Note that turning back to a theory given new data is a kind of induction, since the logical reasoning employed here goes from a concrete example to a set of abstract statements.

Once again, this argument used in science is not necessarily correct. As noted at the beginning of this section, scientists evaluate a theory by testing a prediction which is deduced from a theory. This is because a theory is too abstract to test directly, and we have no choice except testing the concrete prediction. More specifically, scientists deduce a statement such as "if theory X is correct, then we should observe phenomenon Y given environment Z,"

[1] Odden (n.d.) notes that it is not easy to answer "how to decide to abandon the search for predicted entities," that is, how many observations are enough to reach a conclusion. As in his discussion, the answer may depend on a field of science and a type of prediction. As for astronomy, the scope of the scientific field is huge, and thus a large amount of data may be required. On the other hand, if a theoretical prediction is only related to forest creatures in Tokyo, the data can be relatively small. The question with regards to the size of data required for scientific arguments is still an empirical question. Every conclusion potentially has a "101st apple" problem.

and they conclude the validity of theory X by testing the consequent "we should observe Y given environment Z." However, this is known as "affirming the consequent," and the conclusion is not correct from the logical point of view. Consider the case of a well-established statement "If entity A is a penguin, then entity A is a bird." Even if we conclude that entity A is a bird, the entity is not necessarily a penguin. It can be a sparrow, a kiwi, or a puffin. In the same way, the consequent "we should observe Y given environment Z" may be captured by another theory as well. One might wonder why scientists employ this imperfect logic. Fukuzawa (2017: 121) provides an answer to this question. Scientists take it for granted that the validity of a theory should be improved if it is supported multiple times. Due to this flaw in a scientific argument, it is very important to test a theory many times over. The multiple testing of a theory enables us to come closer to the truth with regards to the universe.

2.4 *Summary* This section noted that the goal of science is to construct a single theory that accounts for the entire universe. However, it is very difficult to develop a single theory that accounts for everything in the universe, and thus scientists divide the problem into bits. In the case of linguistics, the goal may be to develop a single theory that models how knowledge in relation to language is memorized and processed in the cognitive system of a human being, and accounts for the reason why a linguistic phenomenon occurs. It was also reviewed that a theory has to satisfy one requirement (i.e., Falsifiability) and would better have three desirable attributes (i.e., Correctness, Comprehensibility, and Aesthetics).

Finally, let us summarize how a scientific theory is developed. This is illustrated in Figure 1. As noted in Section 2.1, a theory contains multiple hypotheses in a systematic manner. In the initial stage of a scientific field, a tentative theory with a bunch of hypotheses is formed to account for a descriptive generalization on the basis of a small set of concrete examples. After this abduction, the theory will be tested and elaborated through a deduction-induction loop on and on. Scientists use the theory to deduce a specific prediction with regards to a phenomenon that is independent of the initial descriptive generalization. Then, they will test the prediction using either an experimental observation or a non-experimental observation. Given the results, they begin another induction. If the result is in line with the prediction, then the theory is supported and stays intact. It may be tested again in future research. If the result contradicts the theory, then the theory needs to be revised or even discarded. That is, the scientist has to stipulate an ad hoc hypothesis that accounts for why the prediction was not supported. This deduction-induction loop lasts forever, because the scientific argument is not perfect. It is necessary to test the theory on and on, so that we can come closer to the truth hidden in the universe.

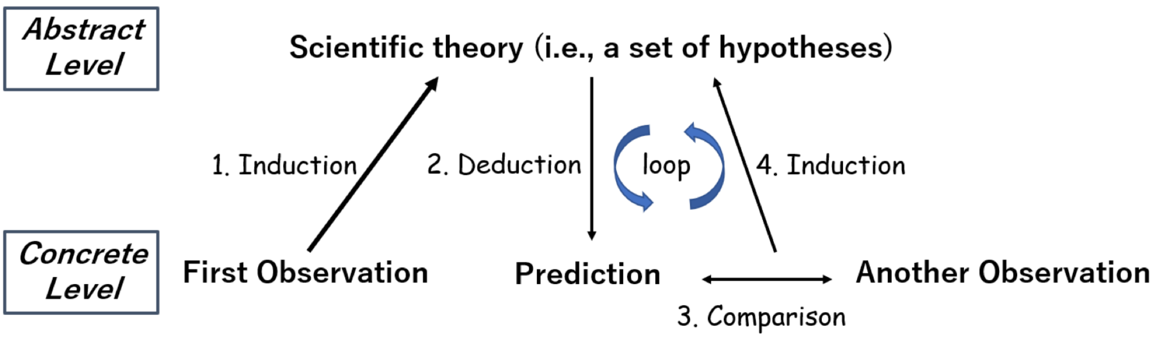

Figure 1. Scientific logics: Deduction-induction loop

3 Scientific argument employed in linguistics: Development of Exemplar Theory

The aim of this section is to review how the scientific logics, illustrated in Figure 1, has been employed in linguistics. In order to understand the logics, we would like to explore how Exemplar Theory has been developed for the last two decades.

Pierrehumbert (2001; 2002) constructed a theory called Exemplar Theory based on observations that were made in previous literature. This theory consists of the following key hypotheses:

(2) Key hypotheses constructing Exemplar Theory[2]
a. Successfully perceived tokens are mentally stored as exemplars (i.e., long-term memories of perceived speech sounds) with random perception noise.
b. Exemplars with similar properties cluster together to form a variety of abstract categories (e.g., phonological, lexical, and sociolinguistic labels).
c. A speaker begins speech production by activating a category that she wants to produce.
d. Based on the activated category, she chooses several exemplars and average the phonetic values to form a production target.
e. A production target is subject to production biases (e.g., production-ease and social norm), and the modified target is produced as a phonetic signal with random production noise.

That is, the cognitive system of a speaker is represented by a tremendous cloud of memorized speech sounds and linguistic labels. These hypotheses are formed to account for a variety of phenomena reported in previous literature: categorical perception, entrenchment, imitation, phonotactic restriction, and word-specific phonetics.

Following this influential induction, many scholars have tested the theory through a deduction-induction loop. By testing phenomena that are independent from the initial induction, the robustness of the theory can be improved. One of the good examples is Hay & Foulkes (2016), which deduced the following statement on the basis of Exemplar Theory:

(3) Prediction 1: Usage effects on selection of a variant
Words likely to be used by younger speakers are produced with innovative variants.

Does this prediction make sense? Unless you are familiar with sociolinguistic literature, it may not make sense. This is because, in addition to the hypotheses of Exemplar Theory, sociolinguistically motivated assumption is required to deduce the prediction. It is necessary to assume that innovative variants are produced by younger speakers more often than by older speakers. Given this assumption, the prediction (3) can be deduced from Exemplar Theory. For example, "computer" is a word likely to be used by young people in New Zealand, while "knitting" is a word likely to be used by older people. A word-medial /t/ is increasingly common to be pronounced as a voiced variant [D] in New Zealand English these days, that is, [D] is an innovative variant while [t] is a conservative variant. This is why "compu<u>t</u>er" is more likely to be produced with [D], and "kni<u>tt</u>ing" is more likely to be with [t]. As a result, a speaker stores more innovative variants in relation to a lexical category "computer," while she stores more conservative variants with regards to a lexical category "knitting." The probability of producing an innovative variant [D] is a function of the number of exemplars encoding the variant, as hypothesized in (2d). Hence, a speaker is more likely to produce an innovative variant [D] when she produces the word "computer" which is likely to be used by younger people. Hay & Foulkes (2016) affirmed the prediction in (3) using corpus research.

Another good deductive example comes from Mendoza-Denton et al. (2003). They put forward the following prediction based on the key hypotheses contained in Exemplar Theory:

(4) Prediction 2: Topic effects on selection of a variant
A speaker is likely to produce a sociolinguistic variant associated with a topic in speech.

Once again, this prediction requires an additional assumption. It is necessary to assume that a topic in speech activates the relevant category. That is, a topic in speech raises the activation of a particular category, with the result that the exemplars belonging to the category are more likely to be chosen for production. Mendoza-Denton et al. (2003) demonstrate that Oprah Winfrey (an American media proprietor) tends to produce more African American English variants when she introduces African American guests. This result can be theoretically interpreted to mean that the topic associated with African American activates the abstract category called "African American English," which activation results in the higher probability of choosing exemplars associated with the category, that is, those mentally stored from speech by African American English speakers.

In this way, Exemplar Theory has been tested through a deduction-induction loop, after it was settled by the first induction. By testing a variety of the predictions, scientists can strengthen the validity of the theory. This is because the theory should capture a wider range of phenomena (i.e., Comprehensiveness). The predictions tested in later research are independent of the descriptive generalizations that the initial induction relies on. The

[2] The original paper has more hypotheses such as weighted activation, granularization and memory decay. Interested readers should refer to the original paper and other important studies (Foulkes & Docherty, 2006; Wedel, 2006; Hay, 2018; Todd et al., 2019).

deduction usually requires additional assumptions besides the key hypotheses of the theory. Scientists always make predictions on the basis of two or more statements including some key hypotheses of the theory and assumptions. It is very important to explicitly state not only the key hypotheses but also tacitly accepted assumptions when we write a paper.

4 Tips to write a scientific paper on linguistics

This final section of this paper gives some tips to help you to write a linguistic paper in a scientific manner. One of the first things I would like to tell is to begin with a particular established theory. Linguistics has a long history, and there are already a variety of influential theories established. Probably, genuine induction is performed only at the very beginning of a scientific field. After the influential induction, what scientists do is to test the established theory via a deduction-induction loop, as demonstrated in Sections 2.3 and 3. Hawkins (1988: 10) also notes that "a new theory is devised that is really an extension of the previous theory." As shown in Section 3, Exemplar Theory has been developed exactly in this way. It stemmed from the influential initial induction by Pierrehumbert (2001; 2002), and has been tested multiple times by many other linguists. A similar theoretical development can be witnessed in generative phonology. Prince & Smolensky (1993/2004) proposed a theory called Optimality Theory, and many other generative linguists tested the theory by exploring the predictions. As the predictions were not found to fit well the observations that were made in the later studies, the theory has been revised with additional hypotheses: Stochastic OT (Boersma & Hayes, 2001); Harmonic Grammar (McCarthy & Pater, 2016); Turbid OT (Tanaka et al., 2017). Recall that the aim of science is to provide a good theory that accounts for the nature of reality. In other words, no theory, no science. As long as you begin with a particular theory, you will absolutely be able to test the theory in a way and contribute to science.

If you make up your mind which theory your research tests, then deduce a specific prediction. This prediction needs to fill a gap in literature, that is, it has to be independent of the predictions tested by previous literature to some extent. There may be at least three ways to deduce a good prediction in linguistics. One way is to test the prediction that was deduced in previous literature using another language. Actually, this extension is quite common in linguistics. For example, Shaw & Kawahara (2019) tested whether the effect of contextual predictability on vowel duration, which is well-attested in English (Aylett & Turk, 2006), extends to Japanese. The second way is to test the prediction that was deduced in previous literature using another similar phenomenon. For instance, Hashimoto (2019) tests a prediction about topic-oriented variation with regards to production of a loanword. This prediction was originally deduced and tested by Mendoza-Denton et al. (2003) in relation to production of a native word. The third way is to deduce and test a brand-new prediction. This type of deduction was reviewed in Section 3, and it may require huge amount of experience and knowledge. In order to take this approach, it might be necessary to study other fields of science such as biology, psychology, and sociology, and import some theoretical concepts or assumptions.

Based on a good deduction, the prediction will be tested using either an experimental observation or a non-experimental observation. It is beyond the scope of this short article how the research methods can be designed, but refer to Di Paolo & Yaeger-Dror (2011) and Drager (2017). There are a variety of scientific methods to test a theoretical prediction.

At the end of this short article, I would like to quote "It's better to be wrong than to be trivial. Ideas that are interesting but wrong are the main engine of progress in this field" from McCarthy (2008: 165). It is always important to try something new, and the failure always grows us. I hope that the discussion in the current article will be helpful for a future linguist:)

References

Aylett, Matthew, & Turk, Alice. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *Journal of Acoustic Society of America*, 119, 3048-3058.

Boersma, Paul, & Hayes, Bruce. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, 32, 45-86.

Chomsky, Noam, & Halle, Morris. (1968). *The sound pattern of English*. New York: Harper and Row.

Clements, George N., & Hume, Elizabeth. V. (1995). The internal organization of speech sounds. In John A. Goldsmith (Ed.) *The Handbook of Phonological Theory* (pp. 245-306). Oxford: Blackwell.

Di Paolo, Marianna, & Yaeger-Dror, Malcah. (2011). *Sociophonetics: A student's guidebook*. New York: Routledge.

Drager, Katie. (2017). *Experimental research methods in sociolinguistics*. Bloomsbury Academic.

Foulkes, Paul, & Docherty, Gerard. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, 34, 409-438.

Fukuzawa, Kazuyoshi. (2017). *Ronriteki Shiko (Logical thinking)*. Tokyo: SB Creative.

Hashimoto, Daiki. (2019). Sociolinguistic effects on loanword phonology: Topic in speech and cultural image. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 10 (1): 11, 1-34.

Hawkins, Stephen. (1988). *A brief history of time*. London: Transworld Publishers.

Hay, Jennifer. (2018). Sociophonetics : The role of words, the role of context, and the role of words in context. *Topics in Cognitive Science*, 10, 1-11.

Hay, Jennifer, & Foulkes, Paul. (2016). The evolution of medial /t/ over real and remembered time. *Language*, 92, 298-330.

McCarthy, John. (2008). *Doing Optimality Theory: Applying theory to data*. Oxford: Blackwell.

McCarthy, John, & Pater, Joe. (2016). *Harmonic grammar and harmonic serialism*. Sheffield: Equinox.

Mendoza-Denton, Norma, Hay, Jennifer, Jannedy, Stefanie. (2003). Probabilistic sociolinguistics: Beyond variable rules. In Rens Bob, Jennifer Hay, Stefanie Jannedy (Eds.) *Probabilistic Linguistics* (pp. 99-138). Cambridge: MIT Press.

Odden, David. (n.d.). Choosing theories is like choosing socks. http://www.ling.ohio- state.edu/~odden.1/ WhatIsTheory.pdf

Okasha, Samir. (2016). *Philosophy of Science: A Very Short Introduction*. Oxford: Oxford University Press.

Oxford University Press. (2007). *Shorter Oxford English Dictionary: Sixth Edition*. Oxford: Oxford University Press.

Pierrehumbert, Janet B. (2001). Exemplar dynamics, word frequency, lenition, and contrast, In Joan L. Bybee, & Paul J. Hopper (Eds.) *Frequency Effects and the Emergence of Linguistic Structure* (pp. 135-157). Amsterdam: John Benjamins.

Pierrehumbert, Janet B. (2002). Word-specific phonetics. *Laboratory Phonology*, VI, 101-139.

Popper, Karl. (1959). *The logic of scientific discovery*. New York: Routledge.

Prince, Alan, & Smolensky, Paul. (1993/2004). *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA and Oxford: Blackwell. [Revision of 1993 technical report, Rutgers University Center for Cognitive Science. Available on Rutgers Optimality Archive, ROA-537.]

Shaw, Jason, & Kawahara, Shigeto. (2019). Effects of surprisal and entropy on vowel duration in Japanese. *Language and Speech*, 62, 80-114.

Tanaka, Shin-ichi, Poppe, Clemens, Hashimoto, Daiki. (2017). Containment eradicates opacity and revives OT in parallel: Some consequences of Turbid Optimality Theory. In Shin'ichi Tanaka, Pintér Gábor, Shinji Ogawa, Mikio Giriko, Hajime Takeyasu (Eds.) *On-in Kenkyu no Shintenkai: Kubozono Haruo Kanreki Kinen Ronbunshu (Festschrift for Professor Haruo Kubozono)* (pp. 40-68). Tokyo: Kaitakusha.

Todd, Simon, Pierrehumbert, Janet B., Hay, Jennifer. (2019). Word frequency eﬀects in sound change as a consequence of perceptual asymmetries: An exemplar-based model. *Cognition*, 185, 1-25.

Toulmin, Stephen E. (1964). *The uses of argument*. Cambridge: Cambridge University Press.

Verserk, Maarten J., Hoogland, Jan, van der Stoep, Jan, de Vries, Marc J. (2015). *Philosophy of technology: An introduction for technology and business students*. New York: Routledge.

Washington State University. (2018). Ask Dr. Universe: Why are apples red? https://askdruniverse. wsu.edu/2016/09/12/why-are-apples-red/

Wedel, Andrew. (2006). Exemplar models, evolution and language change. *The Linguistic Review*, 23, 247-274.

Wells, John. (1982). *Accents of English 1: An introduction*. Cambridge: Cambridge University Press.