

# Psycholinguistic methodology in phonological research

## **Table of Contents**

### Introduction

### Resources that are useful for phonological experimentation

### Anthologies

### History

#### Gap between phonology and experimental approaches

#### The birth and development of laboratory phonology

### The issue of productivity

#### General discussion on wug-tests

#### Specific case studies

### Other offline judgment tasks

#### Word-likeness judgment

#### Naturalness judgment experiment on phonological processes

### Speech perception experiments

#### Probing phonological knowledge through speech perception

#### Probing sources of phonological patterns from speech perception experiments

### Neurolinguistic experiments

### Experiments on language acquisition

### Other psycholinguistic tasks

### Works that address task effects

## **Introduction**

The primary data collection strategy deployed by modern theoretical linguistics is to use intuition of native speakers. This methodology has been criticized throughout the history of generative syntax, although this criticism itself has been much debated (see \*Acceptability judgment[obo-9780199772810-0097]\*). In phonology, the situation is slightly different, as the target of the phonological studies seems less vague or unreliable, although upon closer inspection, solely relying on the intuition-based phonological data could be demonstrably problematic. The majority of phonological data comes in two kinds. One is alternation. A sound can change its shape depending on its phonological or morphological environment; for example, the English plural suffix is pronounced as [s] after *cat*, but as [z] after *dog*. The other is phonotactics: languages have restrictions on how the sounds can be arranged; for example, English allows a [tr] cluster but not a [tl] cluster word-initially. At first sight, neither phonological alternations nor phonotactic patterns seems unreliable for theory construction. Indeed, much of phonological research, at least until 1990's, had developed based on non-experimental data. The data are often gathered based on fieldwork research, dictionary searches, or intuitions provided by native speakers. Questions have been raised, however, regarding whether particular alternations or phonotactic patterns are indeed internalized in the

minds of native speakers. One obvious, yet important, alternative hypothesis is that speakers of a particular language remember all the words in their lexicon, so that these patterns are also remembered on an item-by-item basis. In this view, phonological alternations do not need to be modeled as phonological processes, because all the words are stored in the mental lexicon; i.e. English speakers remember how to pronounce both *cats* and *dogs*, without an abstract phonological principle that governs the realization of the plural allomorph. English speakers also know no words that begin with [tl], without necessarily referring to abstract phonological restrictions. One constructive response to this alternative is to test whether the sound patterns under question can be replicated with nonce words, thereby addressing whether the existing patterns generalize to new words; i.e. whether knowledge under question is *generative*. Much of the psycholinguistic research in phonology has thus been focused on how native speakers produce or respond to nonce words. To the extent that observed patterns are replicable with nonce words, another question that arises is whether the patterns observed in nonce words reflect grammatical knowledge or can be modeled via lexical analogy.

## **Resources that are useful for phonological experimentation**

Cohn et al. 2011 is a very usable handbook on experimental phonology, now also known as laboratory phonology. Statistical methodology is essential in conducting and understanding experimental work. Introductory textbooks which are designed specifically for linguists are Baayen 2008 and Johnson 2008, both of which can be read without prior knowledge of statistics. The statistical program that is used in these books is R (R Core Development Team 1993-), which is in fact used by many practicing linguists these days. It is a free, open source software program, for which many additional analytical packages are available. Macmillan & Creelman 2005 is an accessible introduction to Signal Detection Theory, which is essential to understanding and conducting speech perception experiments, or any experiments that have to do with psychophysics. The book also contains discussions on various experimental paradigms. Praat (Boersma & Weenink 1999-) is free software that is extremely useful for phonetic analyses. Praat, like R, is scriptable. Experigen, designed by Becker & Levine 2013, is an online resource that allow researchers to perform phonological experiments online. A website maintained by John Krantz “Psychological research on the net” hosts many online psychological experiments, including linguistic experiments. This website is an efficient way to gather participants online.

Cohn, C. Abigail, Cécile Fougeron & Marie K. Huffman, eds. 2011. *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press.

This book is the most updated and comprehensive handbook on laboratory phonology, as of 2015. It covers a wide range of topics including, overviews of the issues addressed in the current field of laboratory phonology, current models of phonetics and phonology, various experimental methodologies, and statistical standards in the field.

Baayen, Herald R. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

This is an introductory textbook to statistical analyses written for linguists, using R. The book contains a whole chapter on linear-mixed modeling, which is becoming the common practice in experimental linguistics, instead of more traditional ANOVA.

Becker, Michael & Jonathan Levine (2013) Experigen – an online experiment platform. Available \*online[<http://becker.phonologist.org/experigen>]\*.

Experigen is an online platform to perform various types of phonological experiments online.

Boersma, Paul & David Weenink. 1999-. Praat: Doing phonetics by computer. Software. Software available \*online[<http://www.fon.hum.uva.nl/praat/>]\*. UCLA Praat script resources \*website[<http://www.linguistics.ucla.edu/faciliti/facilities/acoustic/praat.html>]\*.

Praat is free software that executes various phonetic analyses, including acoustic analyses and perception experiments. The software comes with a scripting function, which automates many repetitive processes. Many scripts are made available online by a number of people. It can also implement some statistical analyses and learning algorithms in Optimality Theory.

Johnson, Keith. 2008. *Quantitative methods in linguistics*. Malden and Oxford: Blackwell.

This is an excellent introduction to quantitative data analyses for linguists in general. It covers common statistical analysis techniques used in phonetics, psycholinguistics, sociolinguistics, historical linguistics, and even syntax.

Krantz, H. John (no date) Psychological research on the net, available at the following \*URL[<http://psych.hanover.edu/research/exponnet.html>]\*.

This website lists many online experiments. This website is an efficient platform to gather participants for online experimentation.

Macmillan, Neil & Douglas Creelman. 2005. *Detection Theory: A User's Guide, 2nd Edition*. Mahwah: Lawrence Erlbaum Associates Publishers.

This book offers an accessible introduction to Signal Detection Theory, a theory of psychophysics, which allows us to calculate a measure of sensitivity ( $d'$ ), as well as a measure of bias ( $c$ ). It also discusses several experimental paradigms for speech perception experiments. Some basic understanding of statistics is assumed.

R Core Development Team. 1993-. R. Software. Available \*online[<http://www.r-project.org>]\*.

This free software allows us to perform various statistical analyses. Many packages are available for many different statistical, computational, phonetic and

psychophysical analyses. It also comes with a scripting function, which automates repetitive processes, such as resampling and data processing.

## Anthologies

There are some anthologies that collectively address the issue of using phonetic and psycholinguistic experimentation in phonological research. Ohala & Jaeger 1986 and Solé et al. 2007 are famous examples. Academic journals sometime publish a special thematic issue that is specifically devoted to this issue, which include Beckman 1990, Coetzee et al. 2009 and Ohala 1986, all of which include very important papers. The laboratory phonology series—first published as books and then now as a journal—provides important selection of papers on experimental methodology in phonological research.

Beckman, Mary E., ed. 1990. Phonetic representation. *Journal of Phonetics* 18.

This is a special issue of the *Journal of Phonetics* published in 1990. The volume offers several perspectives on the issue of phonetic representations, and how they are related to phonological representations. Many papers published in this volume have been influential in shaping recent thinking about the phonetics-phonology interface.

Coetzee, W. Andries, Rene Kager & Joe Pater, eds. 2009. Phonological models and experimental data. A thematic issue of *Phonology* 26.

This is a thematic issue of the journal *Phonology*, with a recent collection of papers that exploit a variety of approaches to address phonological issues.

Laboratory Phonology series. The journal \*URL[<http://www.labphon.org/home/journal>]\*.

This series started out as collections of papers that grew out of the laboratory phonology conferences (see **\*the birth and growth of laboratory phonology\***). Ten books have been published. A journal version is launched in 2010. It publishes papers that use laboratory methodology to address phonological questions.

Solé, Maria-Josep, Patrice S. Bedder & Manjari Ohala, eds. *Experimental approaches to phonology*. 2007. Oxford: Oxford University Press.

This is a collection of papers on experimental approaches to phonology, published in 2007 as a festschrift for John Ohala. The volume includes papers that take different experimental approaches to phonological issues, including both phonetic-based and psycholinguistic-based approaches.

Ohala, John J., ed. 1986. *Phonology Yearbook 3, A Thematic Issue: The validation of phonological theories*.

This is a special issue of the journal *Phonology Yearbook* (now *Phonology*), which is entirely devoted to the discussion of experimental approaches to phonology. Both Ohala 1986 and McCawley 1986, discussed under **\*Gap between phonology and experimental approaches\***, appear in this volume.

Ohala, John J. & Jeri J. Jaeger., eds. 1986. *Experimental phonology*. Orland: Academic Press.

As declared in its introduction, this book is the first collection of papers that address phonological issues through experimental methodology.

## History

This section reviews the history of the relationship between theoretical phonology and experimental approaches. The first subsection is about the lack (or insufficiency) of communication between theoretical phonology and experimental methodology. The next subsection is about the recent effort to fill the gap between the two approaches.

### **Gap between phonology and experimental approaches**

One question that has often been asked in modern linguistics is whether the use of introspection-based data collection is justified. This debate has particularly been actively discussed in the area of syntax, and we now witness constructive integration of experimental work with theoretical syntax (see **\*Acceptability judgment[obo-9780199772810-0097]\*** and **\*Experimental linguistics[obo9780199772810-0112]\***). The same holds true with phonology. Just as generative syntax's methodology has been criticized by psychologists and cognitive scientists, generative phonology has been criticized by experimentalists. Ohala 1990 describes the situation as "a turf-war" between phonetics and phonology. See also **\*phonetics[obo-9780199772810-0082]\***, **\*phonology[obo-9780199772810-0033]\***, and **\*the phonetics-phonology interface[obo-9780199772810-0168]\*** for related discussion. There has in fact been a gap between phonetics and phonology, although there are several exceptional experimental work even in early eras of generative phonology (e.g. Berko 1958 discussed under **\*wug-test\*** and Perts & Bever 1975 **\*word-likeness judgment\***; see also Berent 2013). The general separation between experimental phonetics and theoretical phonology is partly due to the fact that historically speaking, phonology "grew out of" phonetics, and hence it needed to defend its territory. For this, and possibly other reasons, theoretical phonology did not make much use of experimental methodology. Among other works by the Prague School of Linguistics, Trubetsky 1939/1969 is the seminal work which tried to separate out phonology from phonetics. The foundational work of modern theoretical phonology is Chomsky & Halle 1968, which did not rely much on phonetics in theory construction (except in Chapter 9), let alone experimental techniques. Throughout the history of generative phonology, there have been constant criticisms against this methodology, most notably by John Ohala. One illustrative paper which clearly voiced against the intuition-based approach is Ohala 1986. McCawley 1986 also encourages generative phonologists to expand their empirical scope by deploying experimentation. Berent 2013 discusses how psychologists and phonologists have in fact shared common interests for a long time, but they

did not interact with each other as much as they should. Berent 2013 encourages more extensive communication between theoretical phonology and cognitive psychology.

Berent, Iris. 2013. *The phonological mind*. Cambridge: Cambridge University Press.

This book observes that phonologists and psychologists have addressed similar issues, but that there has not been sufficient communication between the two fields. Berent argues that phonological theories would benefit greatly by paying more attention to the results of psychological work.

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

This book, also now known as SPE, is the foundational work of theoretical phonology. It is mostly "phonetics-free", except that distinctive features are based on phonetic properties and that Chapter 9 introduces the notion of markedness based on phonetic considerations. It also asserts that the target of the linguistic inquiry is "the actual internalized grammar of the speaker-hearer" (p.4).

McCawley, D. James. 1986. Today the world, tomorrow phonology. *Phonology Yearbook* 3: 27-43.

The paper starts with a general critic of the generative grammarians' attitude to *assume* that their theoretical constructs and the data being analyzed are psychological real, without *testing* this assumption. It argues that once phonologists start deploying experimental methodology, its empirical scope will be widened.

Ohala, J. John. 1986. A consumer's guide to evidence in phonology. *Phonology Yearbook* 3: 3-26.

This paper is a provocative article, listing several types of evidence for theory construction in the order of its reliability, in Ohala's opinion. On a scale of 0-10, "surface phonological patterns" is given the score of 1, and "experiments" is given the score of 9.5.

Ohala, J. John. 1990. There is no interface between phonology and phonetics: A personal view. *Journal of Phonetics* 18: 153-171.

The paper provides a historical review of the gap between phonetics and phonology. The paper argues that phonetic and phonological research should be conducted in tandem with each other, and that experimental and theoretical research should be integrated together so much so that phonetics and phonology should not be considered as two separate systems (hence having no interface).

Trubetsky, S. Nikolai. 1939. *Grundzüge der Phonologie*. Göttingen: van der Hoeck & Ruprecht. Translated 1969 by Christiane A. M. Baltaxe as *Principles of phonology*. Berkeley & Los Angeles: University of California Press.

This book is one historical landmark in phonological research, which established phonology as a discipline that is independent of phonetics. It even goes so far as to argue that acoustic and articulatory properties of sounds are only important for the study of phonetics, not for phonology.

### **The birth and development of laboratory phonology**

Although there had always been a few influential experimental works on phonology since 1950's, a clear historical turning point for the relationship between phonetics and phonology was the birth of laboratory phonology. This general research program intends to use experimental techniques to address phonological issues (see also \*the phonetics-phonology interface[9780199772810-0168]\*). The first laboratory phonology conference was held in 1988, and held biannually in subsequent years (see Cohn 2010 and Pierrehumbert et al. 2000 for informative reviews of its history). Beckman & Kingston 1990, an introduction to the first laboratory phonology book, contains a clear declaration of the need for collaboration between phonologists and experimentalists. Since then, we witness the rise of use of phonetic and psycholinguistic experimentation in phonological theorization. Coetzee et al. 2009 provide some anecdotal—admittedly limited, yet very telling—evidence: between 1976-1986, there were 10 doctoral dissertations on phonology written at the University of Massachusetts, Amherst, none of which were experimental; between 1998 and 2008, on the other hand, 8 out of 13 used experimental methodology. As of today, perhaps not everybody would agree that experimentation is *necessary* in phonological research, but few people would disagree to the thesis that it can be *useful*. The association for laboratory phonology, with its associated journal, was launched in 2010. Pierrehumbert et al. 2000 offer a renewed perspective on the field of laboratory phonology as of 2000. Cohn 2010 provides a yet more updated overview of the history of laboratory phonology. Kawahara 2011 offers several concrete topics in which theoretical phonology and laboratory phonology have mutually benefited from each other in the laboratory phonology research. Coetzee 2008 discusses the relationship between phonological theory and experimentation, while reporting three concrete experiments that address the psychological reality of phonotactic restrictions in English.

Beckman, E. Mary & John Kingston. 1990. Introduction. In *Papers in laboratory phonology I: Between the grammar and physics of speech*. Edited by John Kingston & Mary E. Beckman, 1-16. Cambridge: Cambridge University Press.

This paper is an introduction to the first laboratory phonology book. It declares the start of the new research program, laboratory phonology. It discusses the division of labor between phonetics and phonology as two separate fields, and argues that this division of labor could be harmful. The paper encourages collaboration “between phonologists and other speech scientists” (p.5).

Coetzee, W. Andries. 2008. Grammaticality and ungrammaticality in phonology. *Language* 84: 218-257.

It contains general discussion on the role of experimentation in theoretical phonology. It also addresses the question of whether some experimental results should be attributed to grammatical factors or lexical frequency effects (for more on this topic, see **\*Word-likeness judgment\***). The paper reports three experiments on place restrictions in [sCVC] words in English.

Coetzee, W. Andries, René Kager & Joe Pater. 2009. Introduction: Phonological models and experimental data. *Phonology* 26: 1-8

This paper is an introduction to the thematic issue that is devoted to the experimental methodology in phonology. It discusses how and why recent generative phonologists have started to incorporate insights from experimentation. The paper also contains the anecdotal evidence for the growth of laboratory phonology.

Cohn, Abigail. 2010. Laboratory phonology: Past successes and current questions, challenges, and goals. In *Laboratory Phonology* 10. Edited by Cécile Fougeron, Barbara Kühnert, Mariapaola D'Imperio & Nathalie Vallée, 3-29. Berlin & New York: Mouton de Gruyter.

The article provides a historical overview of the laboratory phonology conferences, as well as its development, as of 2010. It offers a summary of the questions that experimental phonology can and should address. The paper contains suggestions for future directions as well as reexaminations of the assumptions of the linguistic theory.

Kawahara, Shigeto. 2011. Experimental approaches in theoretical phonology. In *The Blackwell companion to phonology*. Edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice, 2283–2303. Oxford: Blackwell-Wiley.

This paper provides an overview of the division of labor between phonetics and phonology, and its reunion through the laboratory phonology. It provides an overview of various kinds of experiments deployed to address phonological issues. It also argues that phonological theories have provided insights into what to look for in experimental studies.

Pierrehumbert, Janet, Mary E. Beckman & Robert D. Ladd (2000) Conceptual foundations of phonology as a laboratory science, In *Phonological knowledge*. Edited by Noel Burton-Roberts, Phillip Carr, & Gerard Docherty, 273-303. Oxford University Press, Oxford.

This paper contains a historical overview of the laboratory phonology tradition as of 2000, as well as discussion on the mutual interaction between theoretical phonology and laboratory phonology. The paper also offers a range of topics that have been—and should continue to be—addressed in the laboratory phonology tradition.

## The issue of productivity

This section discusses arguably the most common type of experimentation in phonological research, namely, wug-tests. It starts with the general discussion on wug-tests, followed by overviews of specific case studies.

### General discussion on wug-tests

As stated in the **\*introduction\***, one major issue that is addressed in phonological experiments is the issue of productivity. One major pattern that phonologists analyze is alternations, where sound /A/ becomes [B] between [C] and [D]. Generative phonology sometimes assumes that whenever we observe an alternation of this sort, that alternation is internalized as a part of phonological knowledge. However, rather than assuming this thesis, it is healthier to test whether this alternation under question is productive and psychologically real, using nonce words (see McCawley 1986 and Ohala 1986 discussed under **\*Gap between phonology and experimental approaches\***). If that alternation applies to new words, we have reason to think that that alternation is a part of phonological knowledge, rather than all the existing items being memorized. This test was informally demonstrated by Halle 1978 who argues that English speakers know how to properly pluralize a foreign word like *Bach*, which ends with a non-familiar sound (a voiceless velar fricative [x]): English speakers choose the right allomorph, [s]. In this sense, the use of nonce words has been an inherent part of the generative enterprise since its inception. No large-scale experimentation was integrated into the generative enterprise until recently, however. A more extensive experiment using nonce words was conducted by psychologists, most famously by Berko 1958. Because her experiment uses a now-famous nonce word *wug*, this testing format is now called “wug-test”. In Berko 1958, English-speaking children are presented with a nonce word and asked to inflect that word. In this way, researchers can observe whether the participants apply the phonological alternation under question to nonce words. This methodology has been applied to many other phenomena in many other languages since then. Even if wug-tests show that particular sound patterns apply to nonce words, another important issue that arises is whether the results are due to grammatical knowledge or can be attributed to lexical analogy. This debate is extensively discussed, perhaps more clearly, in the context of word-likeness judgment experiments (see **\*word-likeness judgment\***). Berent 2013 contains a recent extensive review of this debate, and defends the grammatical view, based on a body of evidence from linguistics, psychology, and neuroscience. Van Oostendorp 2012 also contains some discussion on the pros and cons of the use of wug-tests in phonological theorization. Hyman 2007 contains an interesting discussion arguing that fieldwork elicitation, which would be classified as “non-experimental and/or traditional”, needs to be carefully designed, so much as that it should be considered just as experimental—field experimentation is almost like “wug-tests” in the field.

Berent, Iris. 2013. *The phonological mind*. Cambridge: Cambridge University Press.

This book contains a rich summary of previous findings in both experimental linguistics, psychology and neurology, addressing the fundamental question of what

phonological knowledge really consists of. It offers a recent, extensive defense of the view that phonological knowledge is algebraic.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14: 150–177.

This is a classic paper on experimentation using nonce words. The experiment presents English-speaking children with a nonce word like *wug* with a picture. Shown with a picture with two *wug-s*, the children are asked to pluralize the nonce word stimuli.

Halle, Morris. 1978. Knowledge unlearned and untaught: What speakers know about the sounds of their language. In *Linguistic theory and psychological reality*. Edited by Morris Halle, Joan Bresnan & George Miller, 294-303. Cambridge: MIT Press.

The paper contains the classic argument that the nature of linguistic knowledge is generative. Halle argues that this knowledge operates on distinctive features, because English speakers know how to pronounce the plural suffix after a sound that does not exist in English, such as [x], as in *Bachs*.

Hyman, Larry. 2007. Elicitation as experimental phonology: Thlantlang Lai tonology. In *Experimental approaches to phonology*. Edited by Maria-Josep Solé, Patrice S. Beddor & Manjari Ohala (eds), 7-24. Oxford University Press.

In this article, Hyman argues how much solicitation in fieldwork—which is not always considered “experimental”—shares in common with experimental phonology. Elicitation in the field requires careful designing just like wug-tests. He ultimately concludes that elicitation *is* experimental phonology.

Van Oostendorp, Marc. 2012. A consumer guide to phonological evidence. *Nordlyd* 39.2.

This paper reviews several types of evidence used in phonological theorization. It explicitly argues that wug-tests are useful to test the productivity of morphophonological patterns, especially when the number of existing lexical items is limited. It also raises a cautionary remark that the judgment pattern may not purely reflect grammatical knowledge (for more discussion on this problem, see **\*Speech perception experiments\***).

### **Specific case studies**

Wug-tests have been used to test the the productivity of phonological alternations. In some cases, the productivity of some pattern has been questioned. For example, the productivity of velar-softening in English (e.g. *electroni*[k] → *electroni*[s]ity) is challenged by a nonce word experiment reported in Ohala 1974. Later, however, Pierrehumbert 2006 shows that velar softening is productive under certain circumstances (thus implying that denying the productivity of a process with just one experiment could be dangerous). Other examples that question the productivity of certain sound-related patterns include Vance 1991 on several “phonological” changes that occur in Japanese verbal inflectional paradigms, Ohno 2000 on

rendaku in Japanese, and Sanders 2003 on opaque vowel raising in Polish. There is also a more complicated yet informative set of studies, which concerns the absence/presence of post-nasal devoicing. To the extent that phonology is phonetically-natural (see **\*Speech perception experiments\***), languages can turn voiceless obstruents into voiced ones after nasals, but should not show post-nasal *devoicing*. Coetzee & Pretorius 2010 use a wug-test to show that this post-nasal devoicing process is productive in Tswana, which goes counter to that theoretical claim. On the other hand, Gouskova et al. 2011 argue, based on a phonetic experiment, that this devoicing is indeed impossible, as predicted by the phonetically-driven theory of phonology.

Coetzee, W. Andries & Rigardt Pretorius. 2010. Phonetically grounded phonology and sound change: The case of Tswana labial plosives. *Journal of Phonetics* 38: 404-421.

This paper examines an alleged case of post-nasal devoicing in Tswana, which should not be possible under the view that phonology is grounded in phonetics. This paper shows that some speakers do show evidence for productivity of this unnatural phonological pattern, although they also point to evidence that this unnatural system is not stable, and a change is under its way to a more natural system.

Gouskova, Maria, Elizabeth Zsiga & One Tlale. 2011. Grounded constraints and the consonants of Setswana. *Lingua* 121: 2120-2152.

This paper argues that speakers who show post-nasal devoicing actually show devoicing in other contexts as well, thereby casting doubt on the view that a phonetically unnatural process can be productive. They instead defend the view that phonology operates on phonetically-grounded constraints.

Ohala, J. John. 1974. Experimental historical phonology. In *Historical linguistics, vol. 2: Theory and description in phonology*. Edited by John M. Anderson & Charles Jones, 353–389. Amsterdam: North-Holland.

This paper is one of the earliest criticisms against theoretical phonology. It reports a small-scale word-formation task, testing the productivity of velar softening in English, using *-ism*. The results show that velar softening does not often apply to new words. Ohala claims that velar softening in English is not productive.

Ohno, Kazutoshi. 2000. The lexical nature of rendaku in Japanese. In *Japanese/Korean Linguistics 9*. Edited by Mineharu Nakayama & Charles Quinn, 151–164. Stanford: CSLI Publications.

This paper challenges the view that rendaku, a well-studied phenomenon in Japanese, is rule-governed, contrary to what is usually assumed in generative phonology. Using a word-formation task, Ohno argues that the application of rendaku is better modeled via lexical analogy than a phonological rule.

Pierrehumbert, Janet B. 2006. The statistical basis of an unnatural alternation. In *Laboratory phonology 8*. Edited by Louis Goldstein, Douglas H. Whalen & Catherine T. Best, 81–107. Berlin & New York: Mouton de Gruyter.

Building on Ohala 1974, this paper reports a more extensive wug-test of velar softening in English, using the suffix *-ity*. The results show that velar softening applies productively to pseudo-Latinate nonce roots (e.g. *interponic*), but not to non-Latinate-sounding roots (e.g. *mork*).

Sanders, Nathan. 2003. *Opacity and sound change in the Polish lexicon*. Doctoral dissertation, University of California, Santa Cruz.

This dissertation presents a wug experiment to test the productivity of the opaque raising pattern in Polish. In some Polish words, /ɔ/ becomes [u] before underlyingly voiced obstruents, which are devoiced at the surface. Neither of the two speakers show the expected opaque raising pattern with nonce words.

Vance, Timothy. 1991. A new experimental study of Japanese verb morphology. *Journal of Japanese Linguistics* 13: 145-156.

The experiment reported in this paper shows that nonce verbs in Japanese do not exhibit the same alternation patterns as existing verbs. It concludes that Japanese speakers memorize all the conjugated forms, and the sound alternations observed in the verbal inflections are lexical, rather than grammatical.

## Other offline judgment tasks

Wug-tests are not the only (offline) judgment tasks; other formats of experiments are possible, and have been used in phonological research. This section takes up two types of this research tradition, first word-likeness judgment experiments and then naturalness judgment experiments.

### **Word-likeness judgment**

Wug-tests (see **\*wug-test\***) are useful when the phonological patterns under question are alternations. Phonologists are also interested in the patterns of phonotactics, how languages can arrange sounds. For example, English does not allow word-initial consonant clusters like [tl] or [bn]. One important question is whether these gaps are merely accidental or a part of phonological knowledge. One common strategy to address this question is to ask native speakers whether particular nonce words with such gaps are likely to be a part of the native language's lexicon; e.g. "how likely is [tlik] or [blik] going to become a real word in your language?". One concern that is raised, however, is that this task may be better modeled with a lexical analogy to existing words, instead of grammatical constraints. This lexically-oriented alternative analysis was first proposed by Greenberg & Jenkins 1964 and reiterated by some psycholinguistic work, such as Frisch et al. 2000. One clear trend that is found in this sort of experiments is that the judgment patterns are gradient, not a matter of categorical

yes-acceptable vs. not-acceptable, which is hard to reconcile with the standard assumption of generative phonology, and instead seem better modeled with theories based on lexical statistics, for which Daland et al. 2011 offer a comprehensive overview. Taken to its extreme, if phonotactic knowledge can be inferred from the statistical information from the lexicon alone, then grammatical knowledge would be superfluous. Ohala 1986, cited under **\*Gap between phonology and experimental approaches\***, raises this challenge against generative phonology. Bailey & Hahn 2001 and Schademan 2007 admit that both grammatical factors and lexical statistics impact word-likeness judgment patterns, and attempt to address how these two factors interact to shape phonotactic judgment patterns. Kager & Pater 2012 offer an explicit response to the lexicon-only view of phonotactic knowledge. Berent & Shimron 1997 argue for the psychological reality of the OCP—one of the well-known grammatical principles—partly in response to the lexicon-only theory. Pertz & Bever 1974 report a word-likeness experiment in which the participants are asked to offer grammatical judgments for languages that they do not speak. Since these speakers should not know the lexical statistics of the languages they do not know, their results are hard to reconcile with the purely statistical theory of phonotactics. See also Berent 2013, cited under **\*Wug-test\***, and Coetzee 2008, cited under **\*the birth and development of laboratory phonology\*** for extensive reviews and discussion of this “grammar vs. lexicon” issue.

Bailey, M. Todd & Ulrike Hahn. 2001. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 44: 568–591.

This paper presents comprehensive tests of several factors influencing word-likeness judgment patterns, including grammatical phonotactics effects and lexical neighborhood density effects. They show the activity of both grammatical effects as well as lexical effects, the latter being more influential.

Berent, Iris & Joseph Shimron. 1997. The representation of Hebrew words: Evidence from the obligatory contour principle. *Cognition* 64: 39-72.

This paper reports a word-likeness experiment to test the psychological reality of consonant co-occurrence restrictions (a.k.a. OCP) in Modern Hebrew. They argue for the necessity of grammatical constituency structure to account for the obtained results, beyond mere statistical knowledge that can be inferred from the lexicon.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis & Ingrid Norrmann. 2011. Explaining sonority projection effects. *Phonology* 28: 197–234.

This paper provides a comprehensive overview of various lexical statistics theories. The paper also tests how these models account for word-likeness judgment patterns of English speakers, when they are aided by grammatical notions such as syllable structure. They conclude that grammatical knowledge can be projected from lexical statistical patterns.

Frisch, Stephan, Nathan Large & David Pisoni. 2000. Perception of wordlikeness: Effects of segment probability and length on the processing of nonce words. *Journal of Memory and Language* 42: 481–496.

This paper is a representative attempt to derive and explain phonotactic knowledge from lexical statistics. The experiment shows that nonce words that consist of high-frequency subunits are judged more likely to be English words. Words that are longer are considered less likely to be English words.

Greenberg, H. Joseph & James J. Jenkins 1964. Studies in the psychological correlates of the sound system of American English. *Word* 20: 157–177.

This is a seminal work on word-likeness judgment. As with later work, it finds that the native speakers' judgments are gradient (not a matter of yes-possible or not-possible). They argue that the distance from the existing word—typical lexical statistics knowledge—can account for this gradient pattern.

Kager, René & Joe Pater. 2012. Phonotactics as phonology: Knowledge of a complex restriction in Dutch. *Phonology* 29: 81-111.

This paper is an explicit response to the challenge that patterns of phonotactic judgments do not require reference to grammar. Dutch speakers disprefer a syllable containing a long vowel and two consonants, the second of which is non-coronal. They argue that this constraint cannot be learned from the patterns in the lexicon.

Pertz, D. L. & T. G. Bever. 1975. Sensitivity to phonological universals in children and adolescents. *Language* 51: 149–162.

This experiment asked, for various types of words, whether these structures are likely to occur in languages that the participants do not speak; i.e. it is not judgment about their own languages, but speculations about other languages. Those structures that are cross-linguistically rarer tend to be judged to be more unlikely.

Shademan, Shabnam. 2007. *Grammar and analogy in phonotactic well-formedness judgments*. Doctoral dissertation, UCLA.

This dissertation offers a detailed examination of how grammatical factors and lexical factors shape word-likeness judgment patterns. Shademan argues that lexical analogical effects kick in when grammatical factors are not decisive. It also addresses some task effects, such as age and inclusion of real words in the stimuli.

### **Naturalness judgment experiment on phonological processes**

In syntactic studies, sentences are often assigned a scale of grammaticality; for example, "\*" (utterly ungrammatical), "?\*" (ungrammatical, but not entirely bad), "??" (highly unquestionable), "?" (something is odd), and "no mark" (grammatical). This informal use of

a Lickert scale can be deployed in more formal experiments in phonological research (as well as in syntactic research: \*Acceptability Judgments[obo-9780199772810-0097]\*). One advantage of this paradigm is that it can be used for a phonological pattern that is optional. To illustrate, let us take the optional t/d-deletion in English, as the word *west* being pronounced as [wes] without the final [t]. In wug-tests, the participants tend to be "prescriptive" and do not tend to apply such optional processes. Word-likeness judgment tasks would not work very well here either, because both [west] and [wes] are possible renditions of the existing word in English. A naturalness judgment experiment can be used in this context, by asking the participants, for example, "how natural do you think it is to pronounce the word *west* as [wes]?" Kawahara 2011 is an example of the use of a Likert scale for optional devoicing of geminates in Japanese. Gouskova & Roon 2013 present a naturalness judgment experiment for Russian secondary stress in compound formation. Another advantage of this paradigm, compared to wug-tests (see \*wug-test\*), is that the participants can consider all possible forms, and provide their judgments about all of these possible forms. Albright & Hayes 2003 and Tanaka & Yashima 2013 take advantage of this feature of this experimental paradigm.

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90: 119-161.

This work deploys a naturalness rating task for past-tense formation for nonce verb forms in English. It obtains naturalness ratings of different possible inflected forms for each stimulus verb. The ratings obtained in the experiment for each form are used to test the proposed computational language learning model.

Gouskova, Maria & Kevin Roon. 2013. Gradient clash, faithfulness, and sonority sequencing effects in Russian compound stress. *Journal of Laboratory Phonology* 4: 383-434.

This paper offers a naturalness rating study of secondary stress patterns in compound formation in Russian. The results show that various factors affect the naturalness of secondary stress patterns. The overall results improve upon impressionistic observations which were reported previously.

Kawahara, Shigeto. 2011. Japanese loanword devoicing revisited: A rating study. *Natural Language and Linguistic Theory* 29: 705–723.

This paper presents a naturalness judgment experiment on the optional geminate devoicing in Japanese. The experiment supports previously reported intuition-based data, but reveals further complications that go beyond the introspection-based description.

Tanaka, Yu & Jun Yashima. 2013. Deliberate markedness in Japanese hypocoristics. *Proceedings of GLOW in Asia IX*. Edited by Nobu Goto, Koichi Otaki, Atsushi Sato & Kensuke Takita, 283-297.

This experiment examines different hypocoristic forms used in the casual speech of Japanese. They argue that Japanese speakers find hypocoristic forms with marked

structures more natural than those without, a result that goes counter to the prediction of the markedness theory.

## Speech perception experiments

Speech perception experiments have also been used to explore the nature of phonological knowledge. (This article does not deal with articulatory phonetic studies, which can also be relevant to phonological research, since articulation is usually not classified as “psycholinguistic”. See \*phonetics[obo-9780199772810-0082]\*.) In this section, we first review experiments that aim to problem phonological knowledge by examining how phonological knowledge affects speech perception. Then we discuss experiments which seek to explore perceptual underpinning of some phonological alternations.

### Probing phonological knowledge through speech perception

The tasks reviewed in the two preceding sections—wug test, word-likeness judgment, and naturalness rating (see **\*The issue of productivity\*** and **\*Other offline judgment tasks\***)—all involve explicit thinking by the participants of the experiments. One can raise a concern that these sorts of tasks involve various types of additional cognitive processes in addition to grammaticality judgment. Goldrick 2011 provides an informative overview on this issue. One way to more directly test the psychological reality of phonological knowledge is to investigate how alleged phonological patterns affect speech perception (see \*speech perception[obo-9780199772810-0089]\*). Massaro & Cohen 1983 is a classic work in this research tradition, examining how English phonotactic restrictions, in particular the prohibition against word-initial [tl] clusters, affect speech perception patterns of English listeners. One could argue, however, that their results can be explained as lexical frequency effects; perceptual bias against [tl] can be attributed to zero frequency of this cluster in the English lexicon. To address this alternative, Moreton 1999 demonstrates that the prohibition against word-initial [tl]-sequences impacts the speech perception of English listeners, while the prohibition against word-initial [bw] does not, while both [tl] and [bw] are arguably absent from the English lexicon. Similar to Massaro & Cohen 1983, Hallé et al. 1998 show a perceptual bias against [tl] and [dl] clusters for French listeners. Dupoux et al. 1999 demonstrate that Japanese speakers perceptually insert a vowel between two consonants, because their language does not generally allow sequences of two consonants. Berent et al. 2007 show that similar perceptual epenthesis can occur for English listeners for consonant clusters with no or little sonority rise, and especially when there is sonority fall. Moreton & Amano 1999 argue for the psychological reality of the lexical substrata in Japanese, by demonstrating that strata-specific phonotactic restrictions can impact the speech perception of Japanese listeners. Huang & Johnson 2010 report how the knowledge of Mandarin Chinese and English each affects the perception of tonal similarities.

Berent, Iris, Donca Steriade, Tracy Lennertz & Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104: 591–630.

This experiment tests the perceptual illusion effect with English listeners, with three types of unattested clusters; those with insufficient sonority rise (e.g. [bn]); those with

flat sonority (e.g. [bd]); those with falling sonority (e.g. [ld]). Perceptual illusion is more likely in the order of [ld] > [bd] > [bn].

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568–1578.

Japanese does not generally allow consonant clusters (e.g. \*[VbzV]). When Japanese speakers pronounce such clusters, they epenthesize [u] between them. The experiments show that Japanese speakers perceive this epenthetic vowel, even when there is no trace of vowels in the acoustic signal (e.g. [ebzo] is perceived as [ebuzo]).

Goldrick, Matthew. 2011. Utilizing psychological realism to advance phonological theory. In *Handbook of phonological theory (2nd edition)*. Edited by John Goldsmith, Jason Riggle, & Aaln Yu, 631-660. Blackwell.

This article offers an overview of cognitive processes that may influence off-line judgment tasks, such as similarity to existing items, the existence and the lexical frequencies of lexical neighbors, and (mis)perception of the stimuli. It introduces various experimental paradigms that may tap phonological knowledge more directly.

Hallé, A. Pierre, Juan Segui, Uli Frauenfelder & Christine Menuier. 1998. Processing illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24: 592-608.

French does not allow either [dl] or [tl] word-initially. When presented with such stimuli, French speakers perceive them as [gl] or [kl], instantiating a perpetual bias due to phonotactic knowledge. The experiment uses a gating paradigm, which shows that when French speakers hear only the first consonant, they perceive [d] and [t] accurately.

Huang, Tsan & Keith Johnson. 2010. Language specificity in speech perception: Perception of Mandarin tones by native and non-native speakers. *Phonetica* 67: 243–267.

This paper compares the speech perception patterns of Mandarin speakers and English speakers. In a similarity rating experiment, in particular, the effect of their phonological knowledge is clear in that two tones that are neutralized in Mandarin are rated as very similar by Mandarin speakers.

Massaro, D.W. & M. Cohen. 1983. Phonological context in speech perception. *Perception & Psychophysics* 34: 338-348.

In English, word-initially, [r] can occur after [t] but not after [s], whereas [l] can occur after [s] but not after [t] (\*[sr] and \*[tl]). English listeners identify more of the [r]-[l] continuum as [l] after [s], and [r] after [t]; i.e. they are biased against perceiving sound structures that are not allowed in the native language.

Moreton, Elliot. 1999. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84: 55-71.

This experiment tests two kinds of gaps found in the English lexicon: \*[tɫ] and \*[pʷ]. Moreton argues that the former is grammatical, whereas the latter is accidental. The experiment shows that English speakers show perceptual bias against perceiving [tɫ], but not against [pʷ].

Moreton, Elliott & Shigeaki Amano. 1999. Phonotactics in the perception of Japanese vowel length: Evidence for long-distance dependencies. *Proceedings of the 6th European Conference on Speech Communication and Technology*: 2679–2682.

Japanese lexicon consists of sets of vocabularies with different etymologies. There are two distinctive phonological characteristics for Sino-Japanese items: (1) palatalized [rʲ] appears almost exclusively in Sino-Japanese, but (2) word-final long [aa] is not allowed. Listeners show bias against word-final [aa], when they heard [rʲ] elsewhere in the stimuli.

### **Probing sources of phonological patterns from speech perception experiments**

Another productive line of research is to identify the source of phonological patterns in the pattern of speech perception. John Ohala has extensively argued that speech perception non-trivially affects historical sound changes (and hence synchronic phonological patterns). Ohala 1990 for example observes that in CC place assimilation pattern, the dominant cross-linguistic trend is regressive (e.g. /tk/ → [kk]) rather than progressive (e.g. /tk/ → [tt]). He shows that when listeners are presented with C1C2 clusters with short duration, the listeners usually identify it as C2, rather than as C1. Guion 1998 presents a speech perception study which shows that the source of recurrent sound change pattern, \*[kʲ]→ [tʃ], may have its basis in speech perception. Kochetov 2006 shows that cross-linguistic patterns of distribution of secondary palatalization may be based on the perceptibility differences of this contrast in different phonological environments. Kawahara 2006 argues that geminate devoicing found in Japanese loanwords derives from the fact that voicing in geminates is not very perceptible. Kochetov & So 2007 report a set of perception experiments that address the connection between the susceptibility of place of articulation and perceptibility of place contrast. Mielke 2003 reports a cross-linguistic experiment, which shows that deletion of /h/ in Turkish is more likely to occur in the environment where the presence of /h/ is less perceptible. Mielke shows at the same time that, like many studies discussed in **\*Probing phonological knowledge through speech perception\***, speech perception patterns are non-trivially affected by listeners' language background. McGuire & Babel 2012 explore the reason why the interdental fricative /θ/ is typologically rare, often being replaced with /f/; they argue that articulations of /θ/ is highly variable, and is confusable in terms of both visual and audio perception. One general debate concerning the relationship between speech perception and phonological patterns is whether the influence is through diachronic changes, as argued for example by Ohala 1990 and Guion 1998, or directly synchronic, as argued by Kawahara 2006 and others. de Lacy & Kingston 2013 offer a useful summary of this debate, as well as some

arguments for the latter view, while admitting that speech perception patterns do affect diachronic changes in some cases.

de Lacy, Paul & John Kingston. 2013. Synchronic explanation. *Natural Language and Linguistic Theory* 31: 287-355.

This paper contains a review of the debate about whether speech perception patterns affect synchronic phonological patterns directly, or only indirectly through diachronic changes. While the paper admits that there are cases in which the influence is indirect and diachronic, it offers some evidence for the synchronic effect of speech perception on grammar.

Guion, Susan. 1998. The role of perception in the sound change of velar palatalization. *Phonetica* 55: 18–52.

A recurrent historical sound change is for palatalized [kʲ] (and [k] before front vowels, which are usually palatalized too) to become [tʃ]. The paper first shows that [kʲ] and [tʃ] are acoustically very similar. The perception results show that English listeners can often misperceive [kʲ] as [tʃ]. The general aim of the paper is “to replicate a sound change in a laboratory”.

Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82: 536–574.

Japanese geminates devoice to dissimilate from another voiced obstruent. In this devoicing pattern, only geminates devoice, but not singletons. The paper reports an acoustic and perception experiment to show that devoicing of geminates is not perceptually conspicuous, whereas devoicing of singleton is highly noticeable.

Kochetov, Alexei. 2006. Testing licensing by cue: A case of Russian palatalized coronals. *Phonetica* 63: 113–148.

This paper presents acoustic and perception experiments testing the hypothesis that secondary palatalization is more likely to occur in the environment where this contrast is more perceptible. The results largely support the hypothesis, although not all perceptual differences found in the experiment are reflected in phonological patterns.

Kochetov, Alexei & Connie K. So. 2007. Place assimilation and phonetic grounding: A cross-linguistic perceptual study. *Phonology* 24: 397-432.

This paper examines the perceptibility of place contrasts by listeners from four different languages. The perceptibility differences among different place of articulation are more or less consistent across the four languages. The patterns of speech perception, however, only partially reflect cross-linguistic patterns of phonological place assimilation.

McGuire, Grant & Molly Babel. 2012. A cross-modal account for synchronic and diachronic patterns of /f/ and /θ/ in English. *Journal of Laboratory Phonology* 3: 251-272.

/θ/ is typologically rare and often replaced with /f/, both synchronically and diachronically, but not vice versa. This study shows that the articulation of /θ/ is highly variable across speakers. Consequently, the perception of /θ/ is less stable than that of /f/, both in terms of visual and audio perception.

Mielke, Jeff. 2003. The interplay of speech perception and phonology: Experimental evidence from Turkish. *Phonetica* 60: 208-229.

This paper reports a perception experiment which shows that deletion of /h/ in Turkish occurs in the environment in which /h/ is not very perceptible. By studying the perception patterns with different language background, the experiment also shows that the perception patterns are non-trivially affected by the language background.

Ohala, J. John. 1990. The phonetics and phonology of aspects of assimilation. In *Papers in laboratory phonology I: Between the grammar and physics of speech*. Edited by John Kingston & Mary E. Beckman, 258–275. Cambridge: Cambridge University Press.

This paper presents a perception experiment addressing why phonological/historical place assimilation in consonant clusters is dominantly regressive (i.e. /tk/ → [kk]) rather than progressive (i.e. /tk/ → [tt]). When presented with non-homorganic consonant clusters with short duration, listeners dominantly identified them as second consonants rather than first consonants.

## Neurolinguistic experiments

As with other areas of linguistic research, there have been many attempts to tackle phonological knowledge through neurolinguistic studies. Neurological studies have for a long time been examining the patterns of aphasia: damages to particular portions of brains can result in different types of disorders. Gandor 1998 provides a useful overview of how different types of aphasia can result in different types of phonetic and phonological disorders. Going beyond the study of aphasia, due to recent technological developments, there is now a substantial body of work that deploys neuroimaging techniques to address phonological theories. Mohanan et al. 2012 provide a useful summary of the history and several brain-imaging techniques that are currently available. Näätänen et al. 1997 is a pioneering work in this area, which applied mismatch negativity (MMN) paradigm to the study of phonological issues. In the MMN paradigm, listeners are presented with a series of the same sounds; when there is a change in stimulus, the brain shows a particular response. Extending on the experimental finding by Dupoux et al 1999 (discussed under **\*Probing phonological knowledge through speech perception\***), Dehaene-Lambertz et al. 2000 show that Japanese speakers perceptually epenthesize a vowel between two consonant clusters at the very early stage of speech recognition, to the degree that Japanese brains seem to hear the illusionary vowel. Berent et al 2014 examine the brain responses in Broca's area for the different types

of sonority profile violations in English onset clusters, whose behavioral patterns are first explored by Berent et al. 2007 (discussed under **\*Probing phonological knowledge through speech perception\***). Cornell et al. 2011 argue for the Featurally Underspecified Lexicon (FUL: discussed under **\*Other psycholinguistic task\***), using the MMN paradigm. Kobayashi et al. 2014 addressed the question of whether rendaku, a well-known sound pattern in Japanese, is grammatical or lexicalized, by measuring ERP responses. The development of very non-intrusive neuro-imaging devices, such as NIRS and fMRI, have also recently yielded new insights into the phonological development, and addressed the debate of innateness rather directly by examining the speech perception patterns of newborns. Peña et al. 2003 is a well-known study in this type of work, which shows that newborns already show the left hemisphere dominance for speech perception, which is in accordance with the innateness hypothesis asserted by generative linguistics.

Berent, Iris, Hong Pan, Xu Zhao, Jane Epstein, Monica L. Bennett, Vibhas Deshpande, Ravi Teja Seethamraju & Emily Stern. 2014. Language universals engage broca's area. *PLoS ONE* 9(4): e95155.

This paper offers a neurolinguistic exploration of the effects of sonority in consonant clusters found in Berent et al. 2007 (**\*Probing phonological knowledge through speech perception\***). In behavioral patterns, the perceptual epenthesis is more likely in the order of [ld] > [bd] > [bn]. The same hierarchy is also found in the activity of Broca's area, finding the neurological basis of perceptual epenthesis effects.

Cornell, A. Sonia, Aditi Lahiri & Carsten Eulitz. 2011. What you encode is not necessarily what you store: Evidence for sparse feature representations from mismatch negativity. *Brain Research* 1394: 79-89.

This paper uses a MMN paradigm to test the hypothesis that in the mental lexicon, redundant features are underspecified (Lahiri & Reetz 2002 cited under **\*Other psycholinguistic tasks\***). Coronals are underspecified underlyingly, and should not cause conflict with incoming dorsal signals. Dorsals are underlyingly specified, and hence should conflict with incoming coronal signals. This experiment finds larger MMN responses in the latter case, supporting the prediction of the underspecification theory.

Dehaene-Lambertz, Ghislaine, Emmanuel Dupoux & A. Gout. 2000. Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience* 12: 635-647.

This paper reports an ERP-based testing of the perceptual illusion effect found by Dupoux et al. 1999 (**\*Probing phonological knowledge through speech perception\***). It uses a MMN paradigm in which one stimulus was repeated several times, after which either the same or different stimulus is presented. Japanese speakers show little if any ERP responses when the last stimulus is different from the first set of sounds by the presence of an epenthetic vowel.

Gandor, Jackson T. Phonetics and Phonology. 1998. In *Handbook of neurolinguistics*. Edited by Brigitte Stemmer & Harry A. Whitaker, 207-219. San Diego: Academic Press.

This chapter presents a topic-by-topic overview of how each type of phonetic and phonological deficits is affected by several types of aphasia. This chapter can be used as a reference guide on the study of aphasia, especially how the study of aphasia has informed us about how different parts of the brain are related to phonology.

Kobayashi, Yuki, Yoko Sugioka & Takane Ito. 2014. Rendaku (Japanese sequential voicing) as rule application: An ERP study. *NeuroReport* 25: 1296-1301.

This paper presents an ERP-based experiment to address the debate about whether rendaku is grammatical or lexical. Some items do not undergo rendaku; when rendaku overapply to such items, it induces LAN and P600, which are observed as a result of overapplication of regular rules in other languages. The paper concludes that rendaku is a regular phonological process.

Monahan, Phillip, Ellen Lau & William Idsardi. 2013. Computational primitives in phonology and their neural correlates. In *Handbook of biolinguistics*. Edited by Cedric Boeckx & Kleanthes Grohmann, 233-256. Cambridge: Cambridge University Press.

This handbook article offers a useful resource containing the theoretical background, technical development, and overviews of different brain imaging techniques that are currently available. The discussion revolves around the issue of what phonological primitives are.

Näätänen, R., Lehtikoski, A., Lennes, M., Cheour, M., Huottilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J. & Alho, K. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385: 432-434.

This article is arguably the best known article among those that uses neuro-imaging technique to address phonological issues. It shows that the magnitude of mismatch negativity depends on the phonemic structure of the native language, showing that knowledge of our native language can impact our neurological behavior.

Peña, Marcela, Atsushi Maki, Damir Kovacic, Ghislaine Dehaene-Lambertz, Hideaki Koizumi, Furio Bouquet & Jacques Mehler. 2003. Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of National Academic of Science (PNAS)* 100: 11702-11705.

This experiment uses optical topography to test the left hemisphere dominance for speech perception. It shows that the human brain is already biased toward using left hemisphere for processing speech at the time of birth. The left hemisphere dominance is found for native speech, but not when the same speech is played backward.

## Experiments on language acquisition

Patterns of language acquisition offer important insights into grammatical theory. Observing child-specific phonological patterns is a classic methodology to study how children acquire the phonological system of their native language, of which Smith 1973 is perhaps the most well-known example. Recall also that the participants of the wug-test by Berko 1958 are children, not adults (see **\*Wug-tests\***). Research on second language phonology usually focuses on their production or perception, but sometimes they aim to address some particular theoretical principles: Eckman 2004 provide a rather thorough overview of this research tradition on second language phonology. Psycholinguistic research has been conducted on infants even before they started speaking, and there are two famous methodologies in this tradition. One is the high-amplitude sucking task. In this task, a speech stimulus is played in response to each sucking action by the infant; the sucking rate goes down, as the infant gets bored with the stimulus (called “habituation”), and a new stimulus is presented. If the new stimulus is perceived as different from the original stimulus, the sucking rate goes up. The classic study of this paradigm is Eimas 1974. Another experimental paradigm is a head-turn technique, which explores infants’ preference by measuring how long they keep their eyes on an object with their head turned. Kuhl 1979 provides an informative overview of these two experimental paradigms. Jusczyk et al. 2009 use the head-turn preference technique to examine the specific prediction of Optimality Theory, in which markedness constraints dominate faithfulness constraints in the initial state of Universal Grammar. A recent development in non-intrusive brain imaging techniques now allows us to neurologically address the issue of language acquisition as well, of which Peña et al. 2003 is a showcase example (cited under **\*Neurolinguistic experiments\***). Yet another line of approach is artificial language learning. In this paradigm, the participants are exposed to certain linguistic patterns in the practice phase and are later tested how well they learn these patterns. This paradigm is useful in revealing what kinds of patterns are easily learnable by human learners, and what kinds of patterns are not; i.e. where there are any learning biases. Moreton & Pater 2012 provide useful overviews of this experimental paradigm.

Eckman, R. Fred. 2004. From phonemic differences to constraint rankings. *Research on Second Language Phonology* 26: 543-549.

This paper provides a thorough review of the issues and methodology in the research conducted on the second language (L2) phonology. Some example topics include the differences between L1 and L2 acquisition, speech production and perception in L2, and several formal theories of L2 phonology.

Eimas, D. Peter. 1974. Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics* 16: 513-521.

This paper is a classic study using a high-amplitude sucking task. The experiment shows that 2- and 3- month old infants show evidence for categorical perception for speech stimuli, but not for non-speech stimuli.

Jusczyk, W. Peter, Paul Smolensky & Teresa Allocco. 2002. How English-learning respond to markedness and faithfulness constraints. *Language Acquisition* 10: 31-73.

They use a head-turn preference paradigm to examine whether infants show evidence for markedness constraints and faithfulness constraints, as well as the ranking of markedness constraints over faithfulness constraints, which are predicted by Optimality Theory. The results show that infants do show preference for structures that conform to the predictions.

Kuhl, K. Patricia 1979. The perception of speech in early infancy. In *Speech and language: Advances in basic research and practice*. Edited by Norman J. Lass, 17-23. Academic Press: New York.

This is an overview article on the speech perception research on infants. After reviewing basic properties of speech perception, it explains in detail two commonly-deployed paradigms: a high-amplitude sucking task and a head-turn technique.

Moreton, Elliott & Joe Pater. 2012. Structure and substance in artificial-phonology learning. Part 1: Structure, Part II: Substance. *Language and Linguistics Compass* 6: 686–701 and 702–718.

These papers provide overviews of an experimental paradigm in which participants learn a new artificial language. This research methodology generally seeks to examine whether learnability bias exists; i.e. whether some linguistic patterns are more easily learned than others.

Smith, Neil. 1973. *The acquisition of phonology: A case study*. Cambridge: Cambridge University Press.

This is the first book-length documentation of child-specific phonological patterns in English within the tradition of generative phonology. It is based on the phonological development of one child, Amahl. Child-specific phonological rules are posited to explain the differences between child phonology and adult phonology.

## **Other psycholinguistic tasks**

Many other psycholinguistic tasks have been used to probe the nature of phonological knowledge. This section offers a selected catalogue of psycholinguistic tasks that are not covered in other sections of this article. Jeager 1986 deploys the concept formation technique to explore how native speakers of English classify phonemes. In this task, the participants are asked to classify sounds, without being specifically told how. Ghallagher 2014 deploys a self-paced repetition task in which the participants orally repeat the stimuli that are presented to them as auditory stimuli. Tamaoka & Makioka 2009 use a speeded repetition task, using written stimuli and analyzing naming latencies and accuracy rates. These tasks help us reveal what kinds of structures are easily pronounceable and what kinds are not. In a lexical

decision task, the participants are asked to decide as quickly as possible whether each stimulus is an existing word or not. Vitevitch & Luce 1999 report a very influential study using this paradigm. Berent et al. 2001 use a lexical decision task to address the psychological reality of OCP in Hebrew speakers, which was previously explored by a word-likeness task in Berent & Simoron 1997 (discussed under **\*Word-likeness judgment\***). Inagaki et al. 2000 develop a so-called "vocal-motor" task, in which the participants—especially children—are asked to make a counting gesture, as they produced stimulus words. In this particular experiment, the children are asked to move a doll on a sequence of circles. This task is useful in revealing what kinds of rhythmic units are deployed in particular languages, and suited for cases in which the participants are children. A yet another often-used task makes use of priming. Priming is the phenomenon in which when a listener hears or sees a word, phonologically and semantically similar words are also activated, which facilitates the recognition of these similar words in later phases of the experiment. Lahiri & Reetz 2002 is an example using this paradigm, who argue that underlying coronal features are underspecified. Beddor et al. 2013 use eye-tracking technique, which is often used for sentence processing studies, in order to investigate the time course of the perception of coarticulatory nasalization.

Beddor, S. Patricia, Kevin B. McGowan, Julie E. Boland, Andries W. Coetzee & Anthony, Brasher. 2013. The time course of perception of coarticulation. *Journal of the Acoustical Society of America* 133: 2350-2366.

Using eye-tracking methodology, this experiment investigates how the perception of coarticulatory nasalization occurs over time. Listeners hear a nasalized vowel before a nasal consonant, and they are presented with two visual choices: CVNC and CVC. They fixated the CVNC choice earlier, when nasalization in the acoustic stimuli started earlier.

Berent, Iris, Joseph Shimron & Vered Vaknin. 2001. Phonological constraints on reading: Evidence from the Obligatory Contour Principle. *Journal of Memory and Language* 44: 644-655.

This paper reports three experiments that address the effect of OCP in a lexical decision task. The results show that Hebrew speakers are sensitive to location of identical pairs of consonants. They also argue that those structures that are grammatically illicit are processed differently from those that are missing from the lexicon accidentally (see **\*word-likeness judgment\*** for related discussion).

Ghallagher, Gillian. 2014. An identity bias in phonotactics: Evidence from Cochabamba Quechua. *Journal of Laboratory Phonology* 5: 337-378.

This experiment uses a self-paced repetition task, in which Quechua speakers are asked to orally repeat the stimuli. The participants are more accurate in repeating words that contain two identical ejective stops than two non-identical ejective stops, despite the fact that both types are unattested in the language.

Inagaki, Kayoko, Giyoo Hatano & Takashi Otake. 2000. The effect of kana literacy acquisition on the speech segmentation unit used by Japanese young children. *Journal of Experimental Child Psychology* 75: 70–91.

This paper develops a vocal-motor task. Japanese-speaking children are presented with a word, and asked to move a doll along with a sequence of circles, while pronouncing a chunk of the stimulus word. This test is suited to observe what rhythmic units children use for speech segmentation.

Jaeger, Jeri J. 1986. Concept formation as a tool for linguistic research. In *Experimental phonology*. Edited by John J. Ohala & Jeri J. Jaeger, 211-237. Academic Press.

In this task, speakers are instructed to group a set of different sounds, without being told how. In English non-aspirated [k] after [s] is categorized in the same group as word-initial aspirated [k], despite the fact that non-aspirated [k] is phonetically closer with word-initial [g].

Lahiri, Aditi. & Henning Reetz. 2002. Underspecified recognition. In *Papers in laboratory phonology 7*. Edited by Carlos Gussenhoven, Natasha Werner & Toni Rietveld, 637-676. Berlin: Mouton.

This paper uses a priming paradigm to evaluate the theory of Featurally Underspecified Lexicon. Underlying [coronal] is underspecified, and hence incoming [labial] and [dorsal] acoustic signals can be matched, and hence can cause priming. For example, the stimulus *Bahm* activates the German word *Bahn* ‘railway’ which can prime *Zug* ‘train’.

Tamaoka, Katsuo & Shogo Makioka. 2009. Japanese mental syllabary and effects of mora, syllable, bi-mora and word frequencies on Japanese speech production. *Language and Speech* 52: 79–112.

This experiment uses a speeded naming task, in which speakers are asked to read the stimuli as quickly as possible. Both accuracy rates and naming latencies are analyzed. This paper shows that speech production patterns by Japanese speakers are based on syllables rather than on moras.

Vitevitch, Michael S., & Luce, Paul A. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 40: 374-408.

This is a very influential work on the word recognition process, using a series of lexical decision experiments. It argues for a two-level processing model, one lexical and one sublexical. While probabilistic phonotactic effects impact the sublexical processing, lexical neighbors affect the lexical level of processing.

## Works that address task effects

As reviewed throughout this article, many different methodologies are possible to investigate different aspects of phonological knowledge, and researchers often face the question of what kind of experimental paradigms would be best suited to address particular theoretical issues. However, there have not been many works that directly address the issue of task effects in phonological experimentation, although there are a number of notable works. Berent 2008 addresses the issue of the difference and similarity between auditory stimuli and written stimuli. Coetzee 2010 argues that speakers use their grammatical knowledge differently, when the stimuli are presented in isolation or when they are presented in pairs. Daland et al. 2011 and Kawahara 2015 compare a naturalness rating study and a forced-choice study, and both found that the latter experimental paradigm would reveal differences between different items more clearly. The emerging conclusion from these works is that the difference between two particular conditions is more likely to be revealed, if the participants are presented with two items at the same time, each representing the relevant condition (the task format sometimes known as “2 alternative forced choice (2AFC)”). Gerrits & Schouten 2004 is a very informative phonetic study on task effects in speech perception experiments, focusing on what kind of tasks tend to result in categorical perception and what kind of tasks do not. Davidson & Shaw 2012 addressed similar task effects in speech perception experiments, in the context of perceptual illusion patterns. Huang & Johnson 2010 compared an off-line similarity judgment task and a same-different discrimination perception task; these two tasks differ in terms of the degree to which native language background affects the judged similarity of two tonal categories. Wilson et al. 2014 examined the patterns of production of non-native consonant clusters by English speakers in a shadowing task, and argue that non-contrastive acoustic details of the stimuli non-trivially affect the resulting production patterns.

Berent, Iris. 2008. Are phonological representations of printed and spoken language isomorphic? Evidence from the restrictions on unattested onsets. *Journal of Experimental Psychology: Human Perception & Performance* 34: 1288-1304.

This paper compares auditory stimuli and written stimuli, in terms of the perceptual illusion effect in illicit consonant clusters, first tested by Berent et al. 2007 (**\*Probing phonological knowledge through speech perception\***). It shows some similarities and differences between these two types of stimuli.

Coetzee, Andries W. 2010. Grammar is both categorical and gradient. In *Phonological argumentation: Essays on evidence and motivation*. Edited by Steven Parker, 9-42. London: Equinox.

This work compares a word-likeness rating task, in which speakers rate each word individually, and a forced-choice comparative word-likeness task, in which speakers compare the word-likeness of two items. The latter task revealed differences between two items more clearly.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis & Ingrid Norrmann. 2011. Explaining sonority projection effects. *Phonology* 28: 197–234.

This paper reports two studies of word-likeness judgments: a naturalness rating task and a forced-choice task. Although both tasks exhibit similar patterns, the comparison shows that the latter task is better at distinguishing items that are near the bottom end of the scale; i.e. items that are hardly English-like.

Davidson, Lisa & Jason Shaw. 2012. Sources of illusion in consonant cluster perception. *Journal of Phonetics* 40: 234-248.

This paper addresses task effects on speech perception—AX discrimination task and ABX discrimination task—as it relates to perceptual epenthesis effect discovered by Dupoux et al. 1999 (**\*Probing phonological knowledge through speech perception\***). The comparison across the two tasks shows that various factors, including acoustic properties of the stimuli and language background of listeners, affect perceptual illusion patterns.

Gerrits, E. & M.E.H. Schouten. 2004. Categorical perception depends on the discrimination task. *Perception & Psychophysics* 66: 363-376.

This paper reports a set of experiments to show that in some speech discrimination tasks, categorical perception is not observed. The paper also contains an excellent overview of different speech perception paradigms.

Huang, Tsan & Keith Johnson. 2010. Language specificity in speech perception: Perception of Mandarin tones by native and non-native speakers. *Phonetica* 67: 243–267.

This paper compares a similarity rating experiment and an AB discrimination task, using different pairs of Mandarin tones for Mandarin speakers and English speakers. The effect of native language background was more prominent in the similarity rating task than in the AB discrimination task.

Kawahara, Shigeto. 2015. Comparing a forced-choice wug test and a naturalness rating task: An exploration with rendaku. *Language Sciences* 48: 42-47.

This paper compares two experimental paradigms: a naturalness judgment task and a forced-choice task, using the data from two previous studies of rendaku and Lyman's Law, which blocks rendaku. It shows that the forced-choice task better reveals the effect of Lyman's Law on rendaku than the naturalness rating.

Wilson, Colin, Lisa Davidson & Sean Martin. 2014. Effects of acoustic-phonetic detail on cross-language speech production. *Journal of Memory and Language* 77: 1-24.

When English speakers pronounce non-native clusters (e.g. /zd/), they tend to insert a vocalic interval between the two consonants. Using a shadowing task, this study reveals that such patterns of the production of non-native clusters are non-trivially affected by non-contrastive phonetic details of the auditory prompts.