DOI: http://dx.doi.org/10.18778/8142-988-7.06

Perception and the nature of the phoneme[1]

Craig Callender

Georgia College, USA

Abstract

Traditional interpretations of the phoneme have viewed it either in terms of physical properties (Jones, 1944b), psychological reality (Sapir, 1933, cf. Jaeger, 1980), or function, typically to serve in opposition to other phonemes within a phonological system (Saussure, 1915, Trubetzkoy 1939, Penzl, 1971). More recently, some phonologists have questioned the value of phonemes to phonological description in a post-generative world (Goldsmith, 1999).

I argue that Sapir and Trubetzkoy's ideas about the psychological reality of phonemes and the role of contrastive oppositions in sound systems are as relevant as ever, a claim justified by research in phonetics, which has demonstrated the importance of perception in some types of phonological change (Ohala, 1993, cf. Kuhl, 1991 and Sendlmeier, 2000).

This chapter considers four examples of phonological contrast and change that may have been rooted in perception. First, fortis/lenis and geminate/singleton contrasts may constitute phonological oppositions when their members are perceived as different (Lisker, 1957, cf. Penzl, 1974). Second, perceptual ambiguity may have played a role in the lack of affricates post-vocalically for old short stops in Old High German texts (Callender, 2017). Third, perception may be relevant to understanding the English Great Vowel Shift. Liberman (1995) argued that the GVS had no beginning, in that there was always some degree of allophonic variation in vowels. To extend his analysis, I argue that it is the perception of new vowels that may have triggered the shift. Finally, I suggest that perceptual salience may be responsible for the maintenance of /ai/ before voiceless consonants in southern US English, where it is often monophthongized in other phonological environments. As each of the changes discussed is rooted in the perception of new sounds, phonological oppositions and psychological reality remain relevant to our understanding of phonemes.

Key words: Perception, psychological basis of phonemes, phonological oppositions

[1] I extend my thanks to the participants of the conference on Categories and Units in Language and Linguistics, held in April 2018 in Wałbrzych, Poland, for their helpful comments and suggestions. I also thank Meg Campbell Sloan for her help in conducting the interviews described in section 6. Mistakes and omissions are my own.

1. Introduction

This chapter addresses two questions; first, how might perception inform our understanding of the nature of the phoneme? To this end, it also considers how phonetic studies might shape our analysis of the role of perception. Second, what would the implications be for language change?

I will argue that Sapir and Trubetzkoy's ideas about the psychological reality of phonemes and the role of oppositions within the language system are still key to our understanding of distinctive units in language. The rest of the chapter is structured as follows: sections 2 and 3 cover psychological and physical interpretations of the phoneme, respectively, while section 4 deals with functional interpretations. In section 5, I discuss the role of perception to our understanding of the nature of the phoneme, paying particular attention to the results of several scholars' phonetic experiments. I also consider perception as a factor in fortis/lenis and geminate/singleton contrasts. Section 6 covers the relevance of perception to three phonological events, namely the High German consonant shift, the English Great Vowel Shift, and monophthongization in southern US English. I will claim that perception is ultimately responsible for the origin and maintenance of phonological contrasts, and indeed the phonological system of a language itself.

2. Psychological Interpretations

Baudouin de Courtenay is generally credited with introducing the notion of the phoneme (and with coining the term). He defined it as "eine einheitliche, der phonetischen Welt angehörende Vorstellung, welche mittelst psychischer Verschmelzung der durch die Aussprache eines und desselben Lautes erhaltenen Eindrücke in der Seele entsteht = psychischer Aequivalent des Sprachlautes" [a unified concept, belonging to the world of phonetics, which originates by means of the psychological melding of the impressions resulting from the pronunciation of one and the same speech sound – psychological equivalent of a speech sound] (Baudouin de Courtenay, 1895, p. 9). Jones (1957, pp. 187–188) however, noted that the idea of the phoneme predates Baudouin; Korean king Se-Jong, for example, developed an alphabet for Korean in about 1450 AD, and represented two labial sounds (slightly aspirated [p] and [b]), which were thought to be non-distinctive in the language, by one grapheme. Thus, Jones reasoned, Se-Jong must have been aware of the 'sameness' of the two sounds, i.e., that they belonged to one phoneme (1957, p. 188). Jones argued furthermore that Henry Sweet may

have independently discovered the phoneme at about the same time as Baudouin, as he referred to broad and narrow phonetic transcription (1957, p. 189).

Sapir is perhaps the best-known proponent of psychological interpretations of the phoneme. He argued that it was not physical sounds that were relevant to a language system, but rather "significant entities" (1933, p. 46). Sapir recounted that Tony, a native speaker of Southern Paiute, a language indigenous to southwestern Utah and northwestern Arizona, when assigned the task of dividing the word *pa:*ßa 'at the water' into syllables, produced *pa:pa*. In Southern Paiute, stems affect stop consonants that follow them in various ways. *Pa* is a spirantizing stem, and Tony, who as a native speaker had a tacit awareness of the rule of spirantization, reproduced a stop in the second syllable when given time to think about dividing the word, even though **pa* does not exist outside of compounds. Thus, Sapir argued, "the ßa of speech behavior, as a self-contained syllabic entity without immediately preceding syllable, is actually felt as a phonological *pa*" (1933, p. 50). Similarly, John Whitney, his interpreter for Sarcee, an Athabaskan language of Alberta, Canada, when questioned about whether *dini* 'this one' and *dini* 'it makes a sound' were homonymous, said that *dini* 'it makes a sound' ended in /t/. Whitney's judgment may be explained by the fact that Sarcee has two final vowel types: simple vowels, and vowels with latent consonants. The second type historically were followed by consonants, which no longer exist in the absolute forms of the words, but are present in suffixes and sandhi phenomena. Thus, if the suffix *–i* 'the one who…' is added to *dini* 'it makes a sound', the resulting form is *dinit'i* (Sapir, 1933, p. 53). Whitney's judgment was thus phonetically inaccurate, yet demonstrated an awareness of the phonemic difference between the two sounds, according to Sapir (see Twaddell, 1935, discussed in section 4, for counter arguments).

As a student of Sapir, Swadesh apparently also thought of phonemes in psychological terms. He defined them as "percepts to the native speakers of the given language who ordinarily hear speech entirely in terms of these percepts" (1934, p. 118). Phonemes could be delineated by an acceptability judgment, namely "pronouncing a word with some modification in one of the phonemes. If the modification cannot be perceived by a native, it is within the range of normal deviation. If the modification seems to trouble the native, it is an extreme deviation from the norm, a distortion. If the native definitely hears some other word or feels that one has the word wrong, one may conclude that the modification has amounted to the substitution of one phoneme for another" (1934, p. 124). This would seem to anticipate Ohala's ideas about perceptual normalization and phonemic change. So while Ohala's ideas may have been groundbreaking, others were likely on a similar path as early as the 1930s.

3. Physical Interpretations

A number of scholars have rejected psychological interpretations of the phoneme, arguing that it should be understood in purely physical, acoustic terms. Martinet, for example, thought that a phoneme could be understood as a range of sounds surrounding its center of gravity (Martinet, 1952, pp. 4–5). The center of gravity could be thought of as an ideal realization of the phoneme, similar perhaps (although solely in an acoustic, not perceptual, sense) to later notions of phonemic prototypes (see Kuhl, 1991 and Sendlmeier, 2000), which will be discussed in section 5.

Jones, a champion of the physical camp, defined the phoneme as "a family of sounds in a given language which are related in character and are used in such a way that no one member ever occurs in a word in the same phonetic context as any other member" (1944b, p. 178, cf. 1957, p. 191). However, he did not completely reject the psychological element of phonemes, at least not by 1957, when he argued that the mentalist and physical conceptions of the phoneme together "formed the foundation upon which a complete theory of the phoneme had to be built" (1957, p. 191). Furthermore, he noted "I find the physical view more easily comprehensible to the ordinary student of languages than any other. At the same time I do not hesitate at times to resort to psychological criteria" (1957, p. 191). Jones further argued that phonemes "distinguish words from one another" (1957, p. 195), hinting perhaps at something of an affinity for functionalism. Indeed, his discussion of chronemes in particular overlapped with the ideas of Trubetzkoy, perhaps even Saussure.

Jones argued that, even when chronemes (distinctive units of length) have more than two absolute lengths, the phonological opposition remains one between long and short chronemes. For example, the long vowel chronemes in *bead*, *bean*, and *beat* are all of different lengths, as are the short vowel chronemes in *bid*, *bin*, and *bit*, yet the distinction remains two-way: long vs. short (Jones 1944a, p. 161, cf. Jones, 1957, p. 198). Absolute chroneme length is less relevant than relative length. So, while Jones viewed himself as a proponent of the physical conception of the phoneme, he may have had one foot in the functional camp.

4. Functional Interpretations

One of the more common approaches to describing the phoneme (and a guiding principle of Prague School phonology) is as a functional entity that exists in opposition to other sounds in a language. Functional approaches to phonology

probably owe their origin to Saussure, who defined the phoneme in terms that would also indicate some consideration for the role of perception: “la somme des impressions acoustiques et des mouvements articulatoires, de l’unité entendue et de l’unité parlée, l’une conditionnnant l’autre” [the sum of the acoustic impressions and articulatory movements, of an understood unity and a spoken unity, one conditioning the other] (Saussure, 1915, p. 65). For Saussure, language was a system based completely on the opposition of its concrete entities (1915, p. 149, cf. Penzl, 1971, pp. 17–18).

One of the true champions of the functional interpretation was Trubetzkoy, who defined phonemes in several ways, including: “Phonologische Einheiten, die sich vom Standpunkt der betreffenden Sprache nicht in noch kürzere aufeinanderfolgende phonologische Einheiten zerlegen lassen” [Phonological units, which, from the standpoint of the language in question, cannot be divided into smaller sequential phonological units] (1939, p. 34). In another definition, he specified the “phonologische Einheiten” more explicitly as “die Gesamtheit der phonologisch relevanten Eigenschaften eines Lautgebildes” [the totality of the phonologically relevant characteristics of a sound structure] (1939, p. 35). Thus, despite his status as a functionalist, he might have agreed, to some degree, with Jones’s “family of sounds”. He did not think of phonologically-relevant characteristics in Sapirian terms however, and even directly criticized Baudouin’s definition of the phoneme as the “psychischer Äquivalent des Sprachlautes” on the grounds that every phoneme has several possible realizations, and there can thus be no one-to-one correspondence between a speech sound and a psychological equivalent (1939, p. 38).

It is probably fair to describe Trubezkoy as a functionalist with tendencies that overlapped with the physical, acoustic approach. For example, his rules for determining mono- *vs*. polyphonemic sounds invoked some physical properties; phonemes had to be one syllable, for example, and produced through one unified articulatory gesture (1939, pp. 50–51). Trubetzkoy seemed to allow for some flexibility though, at least when it came to geminates, which he argued were monophonemic (although they could be biphonemic when they occurred at morpheme boundaries, see Trubetzkoy 1938, p. 156), bisyllabic, and bimoraic (1938, pp. 164 and 168). His analysis of geminates therefore violated three of his rules for monophonemes, as they are divided over two syllables, longer than single phonemes, and occur in positions typically occupied by two phonemes (1939, pp. 50–54). For further discussion on the segmental interpretation of geminates, including how Trubetzkoy anticipated autosegmental approaches to their modeling, see Callender (2010).

Although Trubetzkoy did not directly address the role of perception in delineating phonemes or in phonological change, he seemed to be aware of some

of what Ohala would later write about categorical perception. He argued that the opposition t-d could only be interpreted as gradual if it were part of a phonemic system with another dental stop with voicing greater than that of /t/ or lesser than that of /d/ (1939, p. 68).

Twaddell (1935, p. 9, cf. Trubetzkoy, 1939, p. 38) criticized proponents of the mentalist interpretation of the phoneme on the grounds that the mind cannot be studied directly. He also doubted Sapir's conclusions about the psychological reality of phonemes for two reasons. First, he argued that speakers (such as John Whitney and Tony, discussed in section 2) who characterized different sounds as identical may have simply failed to make distinctions that trained phoneticians would have. Second, Sapir's informants may have been demonstrating unconscious knowledge not of phonemes, but of morphemes (Twaddell 1935, p. 13).

Twaddell did not hold back his criticism of those in the physical camp either; he was one of the first scholars to address the non-uniqueness of phonemes in certain phonological environments. For example, /p/, /t/, and /k/ overlap to a degree with their lenis counterparts after /s/, as they are voiceless, yet unaspirated. For Twaddell, there was no reason to prefer one series of phonemes over the other in that position (1935, p. 41). Schane (1968, pp. 711–713) argued that sounds could simply be left unspecified in non-contrastive positions, echoing Trubetzkoy's notion of the archiphoneme, which he defined as "die Gesamtheit der distinktiven Eigenschaften…, die zwei Phonemen gemeinsam sind… Wenn im Deutschen der eindimensionale Gegensatz d-t im Auslaute aufgehoben ist, so ist das Oppositionsglied, welches dabei in der Aufhebungsstellung auftritt, vom phonologischen Standpunkte aus weder eine Media noch eine Tenuis, sondern es ist 'der nicht-nasale dentale Verschlusslaut überhaupt'" [the totality of distinctive characteristics… that are common to two phonemes… When the one-dimensional opposition of d-t is removed word finally in German, then the oppositional member that remains in that position is, from a phonological standpoint, neither a media nor a tenuis, but rather the 'non-nasal dental stop in general'] (1939, p. 71).

Twaddell was perhaps unique among pre-generative linguists in his claim that the phoneme was a fictitious unit (1935, p. 53), yet there are clear indications that his thinking on the topic overlapped with the Prague School linguists. He described /p/, for example, as "the sum of *all* those phonological differences which correspond to a bilabial articulation as opposed to alveolar or palate-velar, a voiceless articulation as opposed to voiced, a stop articulation as opposed to fricative" (1935, p. 48). Further, he noted, "We combine the stops of 'pin, sopping, nap' as corresponding to a single phoneme, not because of any constant quantitative characteristic, but because of a constant qualitative relation to the stops of 'bin, sobbing, nab' (among other relations)" (1935, p. 57).

5. Perception

If Saussure, Trubetzkoy, and others are correct that phonemes serve as contrastive elements in a phonological system, then it is worth considering the role of perception to our understanding of them. In section 5.1., I discuss Ohala's (1987, pp. 216–217, cf. 1993) ideas on the role of perception to certain types of language change.

5.1. Categorical Perception

Ohala noted that the allophone inventory associated with a phoneme cannot account for all possible micro-variations in its pronunciation. I sometimes demonstrate this point in my general linguistics classes with the following thought experiment: let's suppose that we were arbitrarily to assign 50 options for voice onset time, constriction duration, and aspiration duration (leaving out formants for the moment, and bearing in mind that there could actually be far more options, as we are dealing in milliseconds) for a stop consonant. 50 x 50 x 50 yields 125,000 possible sounds for a single consonant, which of course would leave us with an impossibly unwieldy phonological system. Therefore, as Ohala explained, we order similar enough sounds into categories (phonemes). As Kuhl (1991, cf. Sendlmeier, 2000, pp. 113–116) argued, we may think in terms of phonemic prototypes, which function as perceptual magnets, pulling in sounds that are similar enough to be perceived as the same phoneme. The breakdown of this perceptual normalization may lead to phonological change (Ohala, 1993, p. 239).

5.2. Phonetic Experiments

Jaeger (1980) tested the notion that speakers group allophones together with their associated phonemes through two experiments. In the first, subjects were conditioned to respond to words containing [k^h] via a low-voltage electrical shock to the finger, which was applied for some, but not all, words containing the sound. Her subjects' Galvanic skin responses to a series of words presented auditorily were then measured (Jaeger, 1980, pp. 235–238). Responses for words containing both aspirated /k/ and non-aspirated /k/ (in /sk-/ clusters) were higher than for words without /k/. Thus, Jaeger argued, subjects had generalized [k^h] to a phonological position without aspiration; their responses to both allophones were the same (1980, pp. 242–243). In the second experiment, subjects were tasked to press a button to indicate the presence or absence of [k^h]. Jaeger found that subjects once again grouped aspirated and unaspirated /k/ together (1980, pp. 246–250). In other words, their perceptions of the two sounds were the same.

In an attempt to understand the nature of the fortis/lenis distinction, Reed and Wang (1961) conducted a perceptual experiment that involved tape splicing to remove the /s/ before fortis stops in words such as *spy*, *sty*, and *sky*. They found that most subjects identified the resulting word onset as a lenis sound; they identified *(s)cab* as *gab* more often than *cab* (1961, pp. 79–80). (Recall Twaddell's concerns about phonemic overlapping, discussed in section 3). They concluded that aspiration was a more relevant cue than voicing for the fortis/lenis distinction (1961, p. 81, cf. Lotz, Abramson, Gerstman, Ingemann, & Nemser, 1960, pp. 71–72), yet they did not consider that there may be different cues to the distinction word initially than after /s/ (see Fink, 1974, p. 153).

Lotz et al. (1960) had similar findings, but noted that perception could vary according to an informant's native language. Native speakers of American English, where the contrast between fortis and lenis stops is neutralized after /s/ (as they are unaspirated and voiceless in that position), consistently heard lenis stops there. By contrast, native speakers of Puerto Rican Spanish, which has distinctive stop voicing but not aspiration, tended to identify stops following /s/ as fortis (1960, p. 74), as did native Hungarian (1960, p. 75) and Thai (which has a three-way stop contrast: voiceless aspirated, voiceless unaspirated, and voiced) speakers (1960, pp. 75–76).

Reed and Wang and Lotz et al. did not consider constriction duration, which is likely salient in fortis/lenis contrasts. Lisker, having conducted a tape-splicing experiment that allowed him to manipulate constriction duration of medial stop consonants, reported some ambiguity in informants' perception of *rupee* in the 70 to 80 ms range. Below 70 ms, subjects reported hearing *ruby*, but for durations greater than 80 ms, they reported *rupee*. When subjects were presented with the word *ruby*, they exhibited some perceptual ambiguity around 105 ms, with most reporting *rupee* for longer and *ruby* for shorter durations. Lisker (1957, pp. 46–47) noted that "the sum of all other cues bearing on the p-b contrast balances the effect of a 30 msec difference in closure duration".

The geminate/singleton contrast is also informative to our analysis of the role of perception in phonological distinctions. As with fortis/lenis distinctions, constriction duration may be the primary perceptual cue to distinguishing geminates and singletons in multiple languages, including German (Dieth & Brunner, 1943, Kraehenmann, 2001), Marathi (Lisker, 1958), Turkish (Hankamer, Lahiri, & Koreman, 1989), and Italian (Giovaradi & Di Benedetto, 1998). There are other possible cues to the fortis/lenis and geminate/singleton contrast, including voice onset time (Klatt, 1991), formant transition (Fujimura, 1971), or both (Stevens and Klatt, 1974). What is key to the role of perception though is the fact that informants consistently report hearing either a fortis or lenis sound (or a geminate or singleton), rather than something in between. Ohala's claims about categorical perception hold true for the contrast.

6. Relevance for Phonological Change

In this section, I consider how perception may play a role in some types of phonological change, beginning with the High German tenues shift, which is traditionally reconstructed (in somewhat simplified form, leaving out geminate affricates) as follows: /p, t, k/ → /p^{h}, t^{h}, k^{h}/ → /pf, ts, kx/ → /ff, ss, xx/ → /f, s, x/ (Braune, 1874). One problem with the traditional reconstruction of the shift is the fact that affricates do not survive post-vocalically where West Germanic short stops had been, and there is little in the historical record to indicate that they had been there. Scheer (2005) therefore argued that the affrication stage for that position should no longer be included in reconstructions. Nevertheless, there is some evidence that affricates may have been present in weak position in the past, namely the *Pariser Gespräche* (Gusmani, 1996), possible further shifting of affricates in dialects (Hoffmann, 1900, Tarral, 1903, Seibt, 1903), OHG affricates after liquids (Braune, 1874), and comparative evidence from Liverpool English (Honeybone, 2001[2]).

In Callender (2017), I argued that affricates may have been present in Old High German in weak position, but that the distinction between affricates and fricatives may not have been perceptually salient for OHG speakers. As such, a scribe could have written a fricative to represent a sound that may have been (phonetically) either a fricative or an affricate. There seems to be some similar ambiguity in Liverpool English (see Honeybone 2001, or Callender 2017 for discussion). Thus, in Old High German, perceptual ambiguity may account for the lack of an expected form (affricates) in the textual record.[3]

The English Great Vowel Shift, which involved raising and diphthongization of long vowels from Middle English, is another phonological event that may have roots in perception. Liberman (1995, pp. 219–222) argued that the shift had no beginning, and that diphthongal variants had always existed for the high long monophthongs. His analysis raises the question why the phonological system did not simply continue along contentedly, with allophonic variation of the old long monophthongs. One possible answer is that, once speakers of late Middle English began to perceive the diphthongs as "different", new phonemes developed and set the chain shift in motion. Perhaps Penzl was indeed on the right track in his claim that all language change is preceded by allophone development (1971, p. 18).

As a third example of the role of perception in phonological change, I would like to consider monophthongization of /ai/ in southern US English, which has

[2] See Callender (2012) for discussion on the reconstruction of the shift's phonetic mechanism.

[3] I do not have space to include a treatment of the OHG forms here, but invite the curious reader to have a look at my discussion in Callender (2017).

been around before voiced consonants for at least a century, but is more recent before voiceless consonants (Anderson, 2002, pp. 86–88, cf. Bernstein, 1993). In an interview that I conducted, together with Meg Campbell Sloan (following Montgomery, 1993), of two native speakers of South Carolina English (a 70-year-old man and a 55-year-old woman), we obtained the following results:

Male (age 70)	**[ai]**	**[a]**	**[aᵊ]**	**Contexts**
Before voiceless C	36 (100%)			36
Before voiced C		12 (48%)	13 (52%)	25
Word final		53 (91%)	5 (8.6%)	58
Female (age 55)	**[ai]**	**[a]**	**[aᵊ]**	**Contexts**
Before voiceless C	25 (100%)			25
Before voiced C		2 (11%)	16 (89%)	18
Word final		60 (98.4%)	1 (1.6%)	61

As the table shows, both speakers consistently maintained a diphthong before voiceless consonants, but exhibited monophthongization or at least diphthongal reduction to [aᵊ] before voiced consonants and word finally. The female speaker's use of the terms *mica mine* and *wildlife* provided an example of her contrasting pronunciation. In both cases, she maintained the diphthong before the voiceless consonants (/k/ and /f/), and reduced it before the voiced consonants (/n/ and /l/). She also offered an acceptability judgment that was telling, namely that speakers from the Piedmont region of South Carolina were "twangy"; as an example, she offered the phrase *white rice*, pronounced with [a] in each word. She seemed unaware that she was, herself, monophthongizing and reducing diphthongs in other phonological environments.

The maintenance of /ai/ before voiceless consonants in southern US English may be due to perceptual salience of the diphthong in that position. Whereas most speakers may not perceive a difference between /ai/ and its reduced or monophthongized forms word finally or before voiced consonants (unless they are paying particular attention to them), they seem to do so before voiceless consonants, where monophthongization is marked as unacceptable. Thus, while perceptual ambiguity may have led to a quick merger of OHG affricates and fricatives in post-vocalic position where West Germanic short fortis stops once stood, perceptual salience of diphthongs before voiceless consonants in southern US English may be responsible for their continued presence there, despite their loss or reduction in other phonological contexts.[4]

[4] Although, as Anderson (2002) has noted, diphthongization before voiceless consonants may be on the ascendance.

7. Conclusions

If we accept the premise of categorical perception in phonology, then we are acknowledging that the phoneme is a contrastive entity, in line with traditional functionalist thought. The central role that perception plays in the development, maintenance, and in some cases, collapse of phonemic contrasts indicates that Sapir was right to emphasize the psychological reality of phonemes. It is the perception of a given group of similar sounds as one phoneme that makes a phonological system viable; without it, the physical/acoustic variation available in the system could quickly become overwhelming.

The breakdown of perceptual normalization may account for some phonemic splits. In Callender (2017) I argued that it may have provided the impetus for positional affrication (word-initially, after /l/ and nasals, and in the place of West Germanic geminates) and spirantization (post-vocalically for old short fortis stops) in the High German tenues shift. While there was likely phonetic variation beforehand, it may have been the perception of affricates and fricatives that triggered the phonological stages of the event. Conversely, perceptual ambiguity may account for the absence of some expected forms, which may explain the quick merger of OHG affricates and fricatives post-vocalically and the general absence of affricates in that position in the OHG textual record. Finally, perceptual salience may contribute to the maintenance of contrasts, such as /ai/ vs /a/ before voiceless consonants in southern US English. A perceptual study of native speakers of this variety is an opportunity for further research.

References

Anderson, B. L. (2002). Dialect leveling and /ai/ monophthongization among African American Detroiters. *Journal of Sociolinguistics, 6*(1), 86–98.

Baudouin de Courtenay, J. (1895). *Versuch einer Theorie phonetischer Alternationen. Ein Capitel aus der Psychophonetik.* Strasbourg: Trübner.

Bernstein, C. (1993). Social causes of phonological variation in Texas. *American Speech, 68*(3), 227–240.

Braune, W. (1874). Zur Kenntnis des Fränkischen und zur hochdeutschen Lautverschiebung. *Beiträge zur Geschichte der deutschen Sprache und Literatur, 1*, 1–56.

Callender, C. (2010). Trubetzkoy, autosegmental phonology, and the segmental status of geminates. In M. Procházka, M. Mala, & P. Šaldová (eds.), *Prague School and Theories of Structure*, (pp. 45–60). Göttingen: V&R Unipress.

Callender, C. (2012). Neogrammarian sound change, lexical diffusion, and the second consonant shift. *North-Western European Language Evolution, 64/65*, 37–75

Callender, C. (2017). Borrowing, incomplete lexical diffusion, and the High German tenues shift. *North-Western European Language Evolution, 70*(1), 57–72.

Cohn, A., Ham, W., & Podesva, R. (1999). The phonetic realization of singleton-geminate contrasts in three languages of Indonesia. *International Congress of Phonetic Sciences, 14*, 587–590.

Fink, R. (1974). Orthography and the perception of stops after s. *Language and Speech, 17*(2), 152–159.

Dieth, E. & Brunner, R. (1943). Die Konsonanten und Geminaten des Schweizerdeutschen experimentell untersucht. *Sache, Ort und Wort, 20*, 737–762.

Fujimura, O. (1971). Remarks on stop consonants – synthesis experiments and acoustic cues. In O. Fujimura, L. L. Hammerich, R. Jakobson, & E. Zwirner (eds.), *Form and substance: Phonetic and linguistic papers presented to Eli Fischer-Jørgensen*, (pp. 221–232). Copenhagen: Akademisk Forlag.

Giovanardi, M. & Di Benedetto, M. (1998). Acoustic analysis of singleton and geminate fricatives in Italian. *Web-SLS, The European Student Journal of Language and Speech, 98*(1), 1–16.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology*. Oxford: Blackwell.

Gusmani, R. (1996). Die hochdeutsche Lautverschiebung in den "Altdeutschen (Pariser) Gesprächen". *Historische Sprachforschung, 109*, 133–143.

Hankamer, J., Lahiri, A., & Koreman, J. (1989). Perception of consonant length: voiceless stops in Turkish and Bengali. *Journal of Phonetics*, *17*, 283–298.

Hoffmann, K. (1900). *Laut- und Flexionslehre der Mundart der Moselgegend von Oberham bis zur Rheinprovinz*. Metz: Buchdruckerei der Lothringer Zeitung.

Honeybone, P. (2001). Lenition inhibition in Liverpool English. *English Language and Linguistics, 5*(2), 213–249.

Jaeger, J. (1980). Testing the psychological reality of phonemes. *Language and Speech, 23*(3), 233–253.

Jones, D. (1944a). Chronemes and tonemes [a contribution to the study of the theory of phonemes]. *Acta Linguistica, 4*(1), 1–10. Reprinted in W. E. Jones & J. Laver (eds.), *Phonetics in Linguistics*, (pp. 159–167). (1973). London: Longman.

Jones, D. (1944b). Some thoughts on the phoneme. *Transactions of the Philological Society*, 119–135. Reprinted in W. E. Jones and J. Laver (eds.), *Phonetics in Linguistics*, (pp. 168–179). (1973). London: Longman.

Jones, D. (1957). The history and meaning of the term "phoneme". Supplement to *Le Maître Phonétique*, *3*(35). Reprinted in W. E. Jones and J. Laver (eds.), *Phonetics in Linguistics*, (pp. 187–204). (1973). London: Longman.

Klatt, D. (1991). Voice onset time, frication and aspiration in word-initial consonant clusters. In R. Baken & R. Daniloff (eds.), *Readings in Clinical Spectography of Speech,* (pp. 226–246). San Diego: Singular and Kay.

Kraehenmann, A. (2001). Swiss German stops: geminates all over the word. *Phonology*, *18*, 109–145.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*(2), 93–107.

Liberman, A. (1995). The beginning and end of the (Great) vowel shifts in late Germanic. In I. Rauch & G. F. Carr (eds.), *Insights in Germanic Linguistics I. Methodology in Transition*, (pp. 203–230). Trends in Linguistics. Studies and Monographs 83. Berlin: de Gruyter.

Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language*, *33*, 42–49.

Lisker, L. (1958). The Tamil occlusives: short vs. long or voiced vs. voiceless? In *Sir Ralph Turner Jubilee Volume, presented on the occasion of his seventieth [seventy-first] birthday[s]*, (pp. 294–301). Madras: Linguistic Society of India.

Lotz, J., Abramson, A. S., Gerstman, L. J., Ingemann, F. I., & Nemser, W. J. (1960). The perception of English stops by speakers of English, Spanish, Hungarian, and Thai: a tape-cutting experiment. *Language and Speech*, *3*(2), 71–77.

Martinet, A. (1952). Function, structure and sound change. *Word*, *8*, 1–23.

Montgomery, M. (1993). Questionnaire for the study of South Carolina English. In A. W. Glowka & D. M. Lance (eds.), *Language Variation in North American English: Research and Teaching*, (pp. 96–104). New York: MLA.

Ohala, J. (1987). Experimental phonology. *Proceedings of the Berkeley Linguistic Society*, *13*, 207–222.

Ohala, J. (1993). The phonetics of sound change. In C. Jones (ed.), *Historical Linguistics: Problems and Perspectives,* (pp. 237–278). London: Longman.

Penzl, H. (1971). *Lautsystem und Lautwandel in den Althochdeutschen Dialekten.* München: Hueber.

Reeds, J. A. & Wang, W. S.-Y. (1961). The perception of stops after *s*. *Phonetica*, *6*(1/2), 78–81.

Sapir, E. (1933). The psychological reality of phonemes. Reprinted in Mandelbaum (ed.), *Selected Writings of Edward Sapir in Language, Culture, and Personality*, (pp. 46–59). (1949). Berkeley: University of California Press.

Saussure, F. de. (1915). *Cours de Linguistique Générale*, publié par C. Bally et A. Sechehaye. Payot: Paris.

Schane, S. A. (1968). On the non-uniqueness of phonological representations. *Language, 44*(4), 709–716.

Scheer, T. (2005). How minimal is phonological change? *Folia Linguistica Historica*, *25*, 69–114.

Seibt, W. (1930). Zur Dialektgeographie der hessischen Bergstraße. *Giessener Beiträge zur deutschen Philologie*, *27*, 5–78.

Sendlmeier, W. (2000). Lautsprachliche Informationsverarbeitung beim Fremdsprachenerwerb. In U. Kauzner (ed.), *Pronunciation and the Adult Learner: Limitations and Possibilities*, (pp. 113–128). Bologna: CLUEB.

Stevens, K. & Klatt, D. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America*, *55*(3), 653–659.

Swadesh, M. 1934. The phonemic principle. *Language*, *10*(2), 117–129.

Tarral, N. (1903). *Laut- und Formenlehre der Mundart des Kantons Falkenberg in Lothr*. Strassbourg: Heitz and Mündel.

Trubetzkoy, N. S. (1938). Die phonologischen Grundlagen der sogennanten “Quantität” in den verschiedenen Sprachen. In *Scritti in onore die Alfredo Trombetti*, (pp. 155–174). Milano: Hoepli.

Trubetzkoy, N. S. (1939). *Grundzüge der Phonologie*. 4th ed., 1967. Göttingen: Vandenhoeck & Ruprecht.

Twaddell, W. F. (1935). On defining the phoneme. *Language*, *11*(1), 5–62.