# Nasal place assimilation and the perceptibility of place contrasts

#### Abstract

Typological studies of place assimilation show that nasal consonants are more likely to assimilate in place than oral stops (Cho, 1990; Jun, 1995, 2004; Mohanan, 1993). Jun argues that this typological asymmetry derives from a difference in the perceptibility of the place contrasts in nasal consonants and in oral stops. Since the place contrasts in nasals are perceptually weaker than the place contrasts in oral stops, speakers are more willing to neutralize the former. However, the evidence for the weaker perceptibility of the place contrasts in nasal consonants in the previous phonetic and psycholinguistic research is mixed (Hura et al., 1992; Mohr & Wang, 1968; Pols, 1983; Winters, 2002). To offer additional experimental findings bearing on this debate, this paper reports two similarity judgment experiments and two identification experiments in noise, which all show the lower perceptibility of the place contrasts in nasal consonants in coda. The results are compatible with Jun's idea that the asymmetry in place assimilation may result from a difference in the perceptibility of place contrasts.

# **1** Introduction

#### 1.1 The issue—why do nasals assimilate in place?

The relationship between phonetic perceptibility and phonological patterns has been often discussed in the literature. One recurring observation is that speakers seem to be more willing to neutralize a contrast that is less perceptible (Huang, 2001; Hura et al., 1992; Kawahara, 2006; Kohler, 1990; Lindblom et al., 1995; Steriade, 1997, 2001, 2008). In other words, speakers tolerate articulatory simplification as long as it is "perceptually inconspicuous"—this type of neutralization is known as "perceptually tolerated articulatory simplification" (Huang, 2001; Johnson, 2003; Kohler, 1990).<sup>1</sup>

Cross-linguistic typological studies of place assimilation show that nasal consonants are more likely to assimilate in place than oral stops (Cho, 1990; Mohanan, 1993; Jun, 1995, 2004). There are no languages in which only oral stops assimilate in place, but there are languages in which only nasal consonants assimilate. Jun (1995, 2004) lists nine languages that instantiate the latter pattern: Brussels Flemish, Diola Fogny, Hindi, Keley-I, Lithuanian, Malayalam, Nchufie, Toba Batak, and Yoruba. The examples from Malayalam in (1)-(2) illustrate this asymmetry (Mohanan 1993: 74). The word-final nasals in (1) assimilate to the following consonant in place; oral stops in (2) on the other hand do not undergo place assimilation.

- (1) Malayalam nasal place assimilation
  - a. [kamalam] 'Kamalam (proper name)'
  - b. [kamalaŋ-karaññu] 'Kamalam cried'
  - c. [kamalan-taticcu] 'Kamalam became fat'
  - d. [kamalañ-caati] 'Kamalam jumped'
- (2) Oral stops do not assimilate in Malayalam
  - a. [aksaram] 'letter'
  - b. [utkarsam] 'progress'
  - c. [saptam] 'eight'

<sup>&</sup>lt;sup>1</sup>See Ohala (1990) (and also Blevins 2004, 2006; Yu 2004 among others) for a related view. Although he emphasizes the role of perceptibility in shaping phonological patterns, in his model, the way in which perception affects phonological patterns is through misperception by listeners, rather than deliberate control of speakers. This paper does not address this general alternative. See Hayes & Steriade (2004), Hura et al. (1992), Martin & Peperkamp (2011) and Steriade (2001) for relevant discussion. See also Boersma (2008) for a proposal that derives the effect of neutralization of less perceptible contrasts as an emergent property of a learning algorithm. This paper focuses on investigating the perceptibility difference of the place contrasts between oral stops and nasals; we do not commit ourselves to any particular theoretical implementation of how to incorporate this perceptibility difference into a phonological grammar. Our choice of using a speaker-oriented description—e.g. "speakers possessing knowledge of perceptibility effects"—should thus be tentative.

Likewise in Hindi, all nasals within a morpheme must be homorganic to the following stop, as in (3), whereas oral stops do not obey this restriction, as in (4) (Jun, 1995; Ohala, 1975, 1983b).

- (3) Hindi nasal-stop clusters
  - a. [p<sup>h</sup>əŋki] 'handful'
  - b. [gend] 'ball'
  - c. [tamba] 'copper'
  - d. [gəŋga] 'Ganges'
- (4) Non-homorganic stop-stop sequences
  - a. [gətka] 'a type of club'
  - b. [gupta] 'last name'

Jun (1995, 2004) argues that the asymmetry between nasals and oral stops comes from the perceptibility of the place contrasts in nasals and oral stops. He argues that the place contrasts in nasals are less perceptible than those in oral stops, and that speakers are thus more willing to neutralize a place contrast in nasals than in oral stops. In other words, nasal place assimilation is "perceptually more tolerable" than oral consonant place assimilation, because the former involves less of a perceptual change. This claim has been echoed by several researchers. Boersma (1998) suggests that "[m]easurements of the spectra...agree with confusion experiments (for Dutch: Pols 1983), and with everyday experience, on the fact that [m] and [n] are acoustically very similar, and [p] and [t] are farther apart. Thus, place information is less distinctive for nasals than it is for plosives (p. 206)" (see also Boersma 2008). Ohala & Ohala (1993) likewise maintain that "[nasal consonants'] place cues are less salient than those for comparable obstruents" (pp. 241-242). Beddor & Evans-Romaine (1995) suggest that "[a]n acoustic-perceptual account of nasal place assimilation might argue that place distinctions are perceptually less salient for nasal consonants than for oral stops" (p.147) and conclude that "place of articulation in syllable-final nasals is not perceptually robust" (p. 164). See also Martin & Peperkamp (2011) for general discussion of this view; for studies on acoustic and perceptual characteristics of nasal place contrasts, see Beddor &

Evans-Romaine (1995), Fujimura (1962), Kurowski & Blumstein (1984), Kurowski & Blumstein (1993), Malécot (1956), Narayan (2008), Repp (1986), and references cited therein.

#### **1.2** Disagreements in previous studies

A question then is whether nasal place contrasts are indeed less perceptible than oral place contrasts. However, the evidence for the lower perceptibility of the place contrasts in nasal consonants in previous phonetic and psycholinguistic studies is mixed.

A similarity judgement task by Mohr & Wang (1968) showed that English speakers judge nasal minimal pairs as more similar to each other than oral consonant minimal pairs. However, in their stimuli, nasal pairs were placed in coda, whereas oral consonant pairs were placed in onset. Since we know independently that place contrasts are generally more perceptible in prevocalic position than in postvocalic position (Benkí, 2003; Fujimura et al., 1978; Jun, 1995, 2004; Ohala, 1990; Steriade, 2001), this result should be taken with caution. Kawahara (2009) presented English listeners with two pairs of pseudo-words (i.e. [ma]-[na] vs. [pa]-[ta]), and asked them in a forced-choice format which pair involved more similar sounds. The result shows that the nasal minimal pair was judged to be more similar than oral consonant minimal pairs. However, this study is based on orthography, and the perceptibility of the place contrasts was tested in onset position, while consonants that undergo assimilation are usually placed in coda position (Beckman, 1998; Jun, 1995, 2004; McCarthy, 2011; Ohala, 1990).

Pols (1983) showed that Dutch speakers perceive the place contrasts more accurately in oral stops than in nasal consonants under different noise conditions, while controlling for the position within words (including word-final position). Hura et al. (1992) performed an identification experiment of various word-final consonants—nasals, voiceless oral stops, and fricatives—in preconsonantal position. They found that nasals showed the highest confusion rate in terms of place, stops next, and fricatives least. Statistically speaking, the difference between nasals and obstruents was significant, but the difference between nasals and oral stops did not reach significance.

Indirect evidence for the lower perceptibility of the place contrasts in nasals has also been presented from the analyses of verbal art, such as rhyming and imperfect puns. It has been known that speakers can pair two non-identical sounds in rhyming (a pattern known as half rhymes) and imperfect puns. When they do so, they prefer to pair two similar sounds (Holtman, 1996; Steriade, 2003; Zwicky, 1976; Zwicky & Zwicky, 1986). Studies of Japanese hip hop rhymes (Kawahara, 2007) and imperfect puns (Kawahara & Shinohara, 2009) show that Japanese speakers are more willing to match nasal consonant pairs than oral consonant pairs. These comparisons in the Japanese data, however, are based on onset position, not in coda position. Nasal pairs are also commonly found in English rock lyrics (Zwicky, 1976) and English imperfect puns (Zwicky & Zwicky, 1986), which appear in coda position (e.g. *mine* vs. *tryin'*). However, no statistical comparisons are made between the frequencies of nasal pairs and those of oral consonant pairs.

To summarize, the studies reviewed so far provide (indirect) evidence that place contrasts are less perceptible in nasals than in oral stops.

On the other hand, there are also a few studies that fail to support the claim that nasal place contrasts are less perceptible than oral stop place contrasts. The second similarity judgment experiment reported in Kawahara (2009), which used auditory stimuli, did not show that nasal place contrasts are less perceptible. However, this study presented the stimulus pairs only once, and therefore conclusions based on these results remain speculative.

Winters (2002) points out that the results of Hura et al. (1992) do point to the right direction, but emphasizes that the difference between nasals and oral stops did not reach statistical significance. He furthermore cites other studies (Singh & Black, 1966; Wang & Fillmore, 1961) that failed to support the weaker percetibility of the place contrasts in nasal consonants. Winters's (2002) own experiments—identification experiments in four listening environments (comfortable listening level, in noise (6dB S/N-ratio and -6dB S/N-ratio) and speech reception threshold (at about 40dB))—did not reveal a difference between nasals and oral stops in terms of the saliency of the place contrasts. The results in fact showed the evidence for higher saliency of oral stops' place contrasts in the speech reception threshold condition, but also showed the evidence for the opposite pattern in the other three conditions.

To summarize, it is not clear from the previous experiments that nasal place contrasts are indeed less perceptible than oral consonant place contrasts, especially in coda. This study offers new pieces of information bearing on disagreement among the previous studies reviewed above. To summarize the research questions, they are (i) do we find a perceptibility difference in place contrasts between nasals and oral stops at all? and (ii) if so, in what environments, and under what conditions? To address these questions, this paper reports two similarity judgment experiments and two identification experiments in noise. The first two similarity judgment experiments test the perceptibility of place contrasts in clear listening environments; Experiment I uses tokens with released stops, and Experiment II uses tokens with significantly weakened releases. The next two experiments are identification experiments in a noisy condition; Experiment III tests the perceptibility of place contrasts in word-final position, and Experiment IV tests it in pre-consonantal position. All experiments support the hypothesis that the place contrasts are less perceptible in nasals than in oral stops. Although the general debate about the perceptibility of place contrasts in pre-contrasts and easily and oral stops needs to be further studied, our results offer a substantial piece of information bearing on this debate.

# 2 Experiment I: Similarity judgment experiment

The first experiment was a similarity judgment study, most directly building on an experimental paradigm used by Mohr & Wang (1968). This study builds to a lesser extent on Greenberg & Jenkins (1964) who compared only voiced stops and voiceless stops (see also Babel & Johnson 2010, Fleischhacker 2001, Huang 2004, Huang & Johnson 2010, Kato et al. 1997 among others for studies using this paradigm to investigate knowledge of perceived similarity). In this experiment, native English listeners were presented with pairs of sounds minimally different in place, and were asked to judge the perceived similarity between the two sounds. The experiment used naturally-produced—but acoustically edited—stimuli. The experiment built upon the previous studies reviewed in section 1.2, but controlled factors that may affect similarity ratings: (i) all the stimuli were placed in post-vocalic position, and (ii) amplitude and pitch were made uniform across the stimuli.

#### 2.1 Method

#### 2.1.1 Stimuli

The three conditions were nasals, voiced stops, and voiceless stops. For each condition, all three combinatorial possibilities of different places were included (i.e. labial vs. coronal, labial vs. dorsal, coronal vs. dorsal). All the stimuli were mono-syllabic and had initial vowel [a]. The target consonants were all placed in coda because place assimilation usually occurs in coda position (Beckman, 1998; Jun, 1995, 2004; McCarthy, 2011; Ohala, 1990). Thus our stimuli consisted of [am-an], [am-aŋ], [ab-ad], [ab-ag], [ad-ag], [ap-at], [ap-ak], and [at-ak].

#### 2.1.2 Recording and acoustic editing

Two female native speakers of English (both from New Jersey) each produced all the stimuli in a sound attenuated booth. One speaker was the second author of this paper. Their speech was recorded through an AT4040 Cardioid Capacitor microphone with a pop filter and amplified through an ART TubeMP microphone pre-amplifier (JVC RX 554V), digitized at 44K with 16 bit quantization level. The stimuli were placed in a frame sentence: "Please say the word X three times." To avoid flapping and reduction of word-final consonants, both speakers released all the word-final consonants. The speakers repeated each token 10 times. Some illustrative spectrograms are shown in Figure 1.



Figure 1: Illustrative waveforms and spectrograms of recorded tokens in Experiment I. Top: [ap]; middle: [ab]; bottom: [am].

The target stimuli were extracted from the frame sentence at zero crossings using Praat (Boersma & Weenink, 1999–2014). To avoid similarity ratings being affected by non-relevant phonetic factors such as differences in amplitude or pitch, the stimuli were re-synthesized with a flat pitch contour at 250Hz and with the peak amplitude of 0.7. Out of 10 repetitions, those that had phonetic distortions (e.g. clipping, heavy creakiness, unintended vowel qualities, nasal bursts) were excluded. After that, four tokens from each speaker were chosen as the stimuli for the listening experiment. Pairs of sounds were created by concatenating two sounds with 500ms silence interval. Within each pair, the target sound with the consonant with the fronter place always appeared as the first member (e.g. [am]-[an], not [an]-[am]).

#### 2.1.3 Procedure

In this experiment, one pair of sounds was presented to our listeners per trial without any orthographic representations of the stimuli. The participants were asked to judge the similarity of each pair using a 5-point-scale: 1. "almost identical", 2. "very similar", 3. "similar", 4. "not so similar", 5. "completely different". Superlab (ver 4.0, Cedrus) on Macintosh computers was used to present the visual and sound stimuli and to record responses. All the participants wore high quality headphones (Sennheiser HD 280 Pro), and registered their responses using an RB-730 response box (Cedrus). The experiment took place in a sound attenuated room.

The experiment started with a practice block with 20 pairs in order for the participants to establish their subjective scale of similarity. These stimuli were unique to the practice block. An experimenter stayed in the listening room during the practice session so that the participants could ask questions after the practice session is over. The main session was organized into two blocks, with a break in-between, each block presenting tokens from one speaker. We blocked the experiments by speaker so that the listeners would not be distracted by individual speech style differences. All pairs of sounds were repeated seven times. Hence for each phonological pair, the listeners judged their similarity 56 times (7 repetitions \* 4 tokens \* 2 speakers). Superlab randomized the orders of the stimuli within each block.

#### 2.1.4 Participants

Twenty-one undergraduate students completed this experiment, but the data from two speakers were not analyzed because they were not native speakers of English. All the participants received extra credit for linguistics courses.

One may argue that English listeners may not be appropriate for this experiment, as English has a prefix that exhibits nasal place assimilation (i.e. *in-*). This alternation in English may make the place contrasts in nasals less distinct, because alternation between two sounds may arguably shrink the perceptual distance between the two (e.g. Hume & Johnson 2003, Huang & Johnson 2010; though see also Steriade 2003). However, using English listeners may not be problematic for three reasons. First, prefixal nasal place assimilation is not without exceptions: *un-* does not undergo place assimilation. Second, the target consonants in the first three experiments are placed in word-final position, and the place contrasts are contrastive in this position for both nasals and oral stops in English. Third, English exhibits assimilation of oral stops across word boundaries as well, as in *ba[g] girl* 'bad girl' (Ellis & Hardcastle, 2002; Nolan, 1992). See Beddor & Evans-Romaine (1995) for relevant discussion.

#### 2.1.5 Statistics

For statistical analyses, a general linear mixed model was run with PLACE (labial vs. coronal, labial vs. dorsal, coronal vs. dorsal) and MANNER (nasal, voiced stops, voiceless stops) as fixed factors (Baayen et al., 2008; Baayen, 2008; Bates, 2005; Jaeger, 2008) using R (R Development Core Team, 1993–2014) and the lme4 package (Bates et al., 2011). (In this paper we use capital letters to represent variable names.) The lme4 package does not compute p-values because the procedure to calculate degrees of freedom has not been known. Therefore, they were calculated by the Markov chain Monte Carlo method using the languageR package (Baayen, 2009).

#### 2.2 Results

Table 1 illustrates the average similarity ratings in Experiment I.

	Nasals	Voiced stops	Voiceless stops
Labial vs. coronal	2.69 (0.08)	3.64 (0.06)	3.98 (0.06)
Labial vs. dorsal	2.49 (0.08)	3.67 (0.06)	4.00 (0.06)
Coronal vs. dorsal	2.57 (0.08)	3.60 (0.06)	4.02 (0.06)
Averages	2.59	3.63	4.00

Table 1: The average similarity ratings in Experiment I (margins of errors for 95% confidence intervals). The lower the value, the more similar the pair.

First of all, the comparison between the three manners of articulation shows that nasal pairs were judged to be most similar to each other; voiced stop pairs were judged to be more similar than voiceless pairs. A general linear mixed model shows that MANNER had a significant impact on similarity ratings (t = 51.06, p < .001), but PLACE did not (t = -1.42, n.s.). A contrast analysis comparing nasals and voiced stops shows that MANNER significantly impacted similarity ratings (t = 36.10, p < .001), and so did PLACE (t = -2.15, p < .05). PLACE was perhaps significant because the labial-coronal pair in the nasal condition has a slightly higher rating than the other two place pairs. More importantly, the significant effect of MANNER shows that nasal pairs were rated more similar than voiced stop pairs. Another contrast analysis compared voiced and voiceless stops, and revealed a difference in MANNER (t = 14.68, p < .001.), but not in PLACE (t = -.03, n.s.). Voiced stop pairs were rated more similar than voiceless stop pairs.

#### 2.3 Discussion

#### 2.3.1 Bearing on the place assimilation asymmetry

The results support the hypothesis that the place contrasts are less salient in nasal pairs than in oral stop pairs. This difference in the perceptibility of the place contrasts may be the reason for the place assimilation asymmetry, as suggested by a number of previous researchers (Beddor & Evans-Romaine, 1995; Boersma, 1998, 2008; Jun, 1995, 2004; Ohala & Ohala, 1993; Steriade, 2001). More generally speaking, this result supports the general principle that speakers are more willing to neutralize less perceptible contrasts (Boersma, 1998; Huang, 2001; Hura et al., 1992; Kawahara, 2006; Kohler, 1990; Lindblom et al., 1995; Steriade, 1997, 2001, 2008).

Winters (2002) raises the hypothesis that "any perceptual differences which exist between nasals and plosives might only emerge under noisy conditions" (p. 12), by comparing previous studies on the perceptibility differences in nasals and oral stops (Hura et al., 1992; Pols, 1983). However, the results above show that the difference between nasals and oral stops does emerge under clear listening environments as well, at least if we use a similarity rating paradigm.

In addition to the difference between nasals and oral stops, we also obtained a difference in similarity ratings between voiced and voiceless consonants. This observation replicates previous similarity judgement studies (Greenberg & Jenkins, 1964; Mohr & Wang, 1968). This difference is also observed in the combinability of consonants in Japanese rap rhymes (Kawahara, 2007). Japanese speakers are more willing to pair voiced stops with mismatched place than voiceless stops with mismatched place in creating rap rhymes.

However, phonologically speaking, we do not know of a language in which only voiced consonants assimilate but voiceless consonants do not; e.g.  $/dg/ \rightarrow [gg]$ , but  $/tk/ \rightarrow [tk]$ . It is possible that further typological research on place assimilation may find a language that instantiates this pattern. To the extent that this pattern is a true gap, it remains as a puzzle why the perceptibility difference between voiceless stops and voiced stops is not reflected in phonological patterns.<sup>2</sup> Diane Bradley (p.c.) raised the possibility that assimilation of voiced consonants is blocked by an independent reason: since voiced geminates face an aerodynamic problem, many languages avoid them (Hayes & Steriade, 2004; Ohala, 1983a; Westbury, 1979; Westbury & Keating, 1986): it is challenging to maintain sufficient transglottal airpressure drop with long obstruent closure while maintaining glottal airflow to sustain voicing. However, while voiced geminates do suffer from this aerodynamic problem, so should unassimilated voiced obstruent clusters, because speakers would need to maintain voicing during long obstruent closure. Note also that place assimilation does not necessarily result in geminates when the targets and triggers differ in manner (e.g. when triggered by fricatives).

<sup>&</sup>lt;sup>2</sup>Schane (1972) proposes an idea which assumes that the perceptual difference between voiced stops and voiceless stops does shape a phonological pattern. He proposes that coda devoicing occurs to enhance place contrasts in coda: voiceless consonants are favored over voiced consonants because the place difference is more salient for voiceless consonants than for voiced consonants. However, coda devoicing can be construed as a case of neutralization of a phonological contrast that is not well perceptible (Steriade, 1997, 2008).

#### **2.3.2** Place effects

Next, some remarks on the patterns of different place pairs are in order. Phonologically speaking, coronals are more likely to undergo place assimilation than labials and dorsals (Cho, 1990; Jun, 1995, 2004; Kochetov & So, 2007; Paradis & Prunet, 1991). If this asymmetry is due to a difference in perceptibility, then this hypothesis predicts that pairs that involve coronals should be judged to be more similar than the labial-dorsal pair: coronals tend to assimilate because their cues are not highly perceptible (Boersma, 1998, 2008; Byrd, 1992; Jun, 1995, 2004; Kochetov & So, 2007). However, this prediction is not borne out in our experiments: the labial-dorsal pairs were not particularly judged to be dissimilar, compared to pairs involving coronal consonants.

We should also bear in mind, however, that in asymmetries in place assimilation, the directionality matters; e.g. it is more likely for coronals to become dorsals than for dorsals to become coronals. On the other hand, the similarity judgment task in the current experiment is symmetric. Since the focus of this paper is the differences in the perceptibility of the place contrasts between different manners of consonants, we will set aside the discussion on differences between place of articulation within each manner.

# **3** Experiment II: Similarity judgment Experiment II

The next experiment tested whether the similarity judgment patterns observed in Experiment I would hold without clear release bursts. As observed in Figure 1, the tokens in Experiment I were clearly released. The role of release bursts in the perception of place contrasts has been well known (Kochetov & So, 2007; Malécot, 1956; Smits et al., 1996; Stevens & Blumstein, 1978; Tekieli & Cullinan, 1979; Winitz et al., 1972). Some authors argue that released consonants resist assimilation (Jun, 2003; Kohler, 1990; McCarthy, 2011; Padgett, 1995), because release bursts provide such a strong cue to the perception of place distinctions. Hura et al. (1992) as well as Winters (2002) used non-released voiceless stops in testing the perceptibility difference between nasals and oral stops. A question arises whether the similarity judgment pattern we observed in Experiment I still holds without clear release bursts.

investigate whether the similarity differences we observed in Experiment I could be due to the clearly released tokens.

#### 3.1 Method

#### 3.1.1 Stimuli

To test whether the perceptual asymmetry between the nasal place contrasts and the oral place contrasts would be observed without release bursts, from the tokens we used in Experiment I we spliced off original releases of voiced and voiceless stops at zero-crossings. Without any bursts, however, the stimuli sounded as if there were no consonants at all. Therefore, we recorded weak releases of one speaker (the second author) for [p, t, k, b, d, g] in the context of  $[\alpha_-]$ . (The other speaker left the lab by the time we ran this experiment, so only the tokens from the second author were used.) We adjusted the average amplitude of the original tokens to 70dB and that of releases to 40dB and concatenated them. To be conservative—i.e. to be biased against the conclusion that the place contrast is less perceptible for nasals—we retained original, clear nasal releases. We also readjusted the average amplitude of nasal tokens to 70dB. Waveforms and spectrograms of edited tokens are shown in Figure 2. As shown in Figure 2, the new releases of the stops are extremely weak—they were there only to signal the presence of word-final consonants.



Figure 2: Illustrative waveforms and spectrograms of edited tokens with weak releases (Experiment II). Top: [ap]; bottom: [ab]. Nasal tokens retained clear release, as shown in Figure 1.

#### **3.1.2** Other aspects

The procedure of Experiment II was identical to Experiment I, except for two aspects. One is that we used speech from only one of the speakers for the reason mentioned above. Second, we included both orders between the two elements in a pair (e.g. [am]-[an] and [an]-[am]). Each pair was thus repeated 56 times (7 repetition \* 4 tokens \* 2 orders). Eighteen native speakers of English participated in this study for extra credit for linguistics classes. No participants who participated in Experiment I participated in this experiment.

#### 3.2 Results

Table 2 shows the average similarity ratings in Experiment II. A general linear mixed model shows that MANNER had a significant impact (t = 30.87, p < .001), but PLACE did not (t = -0.14, n.s.). A contrast analysis comparing nasals and voiced stops shows that nasal pairs were judged to be more similar than voiced stop pairs (t = 13.33, p < .001). PLACE did not turn out to be significant in this analysis (t = -1.63, n.s.). Another contrast analysis compared voiced stops and voiceless stops, and revealed a difference in terms of MANNER (t = 14.31, p < .001.), but not in terms of PLACE (t = -0.34, n.s.).

Table 2:	The	average	similarity	ratings	in	Experiment	Π	(margins	of	errors	for	95%	confide	ence
intervals)	).													

	Nasals	Voiced stops	Voiceless stops
Labial vs. coronal	3.12 (0.06)	3.52 (0.07)	3.72 (0.06)
Labial vs. dorsal	2.68 (0.07)	3.45 (0.08)	3.79 (0.06)
Coronal vs. dorsal	2.79 (0.07)	3.14 (0.07)	3.78 (0.07)
Averages	2.86	3.37	3.76

#### 3.3 Discussion

#### 3.3.1 The perceptibility differences

Compared to Experiment I, the oral consonant pairs were judged to be more similar (voiced stops: 3.63 in Experiment I vs. 3.37 in Experiment II; voiceless stops 4.00 in Experiment I vs. 3.76 in Experiment II), which makes phonetic sense, because Experiment II used oral stops with very weak releases. Nevertheless, these stop pairs without clear releases were judged to be less similar than nasal pairs with clear releases intact.

The results show that even when we replace original clear releases of oral stops with weak ones, the same perceptibility hierarchy of the place contrasts holds: voiceless stops > voiced stops > nasals. Recall that the nasal stimuli retained their original clear releases; i.e. they had advantage in conveying place contrasts, but they were nevertheless judged to be most similar. As with Experiment I, the results support the hypothesis that place of articulation is less perceptually salient in nasals than in oral stops (Jun, 1995, 2004). This perceptual difference holds even when nasals retain their clear releases and oral stops have only very weak releases.

### **4** Experiment III: Identification experiment in noise

The third experiment aimed to verify the perceptibility differences observed in the previous two experiments with an identification task in noise. Hura et al. (1992) ran their identification experiment in a clear listening environment and obtained only 5.2% of misidentification. This low percentage of misidentification may be the reason for why they did not obtain a significant difference between nasals and oral stops. As reviewed in the introduction, a number of other identification experiments in noise have been run in the past, and they showed conflicting results. Pols (1983) found the expected difference between nasals and oral stops, whereas Winters (2002) did not. To add more experimental results bearing on this issue, we ran an identification experiment in noise. What is new in our Experiment III is that it emulates the real communicative situation most closely, by using cocktail party noise to cover the stimuli. There is another motivation for this experiment. The two previous similarity judgment experiments involve an off-line task which involves conscious judgments by listeners. While the results support Jun's (1995; 2004) idea that perceptibility differences underlie the differences in the likelihood to undergo assimilation, it would be ideal to further support this idea by a task that does not involve conscious judgments.

#### 4.1 Method

#### 4.1.1 Target stimuli

The stimulus structure is the same as the previous experiments; all the stimuli were mono-syllabic and had initial vowel [a], consisting of [ap, at, ak, ab, ad, ag, am, an, aŋ]. A female native speaker of English from New Jersey (the second author) pronounced the stimuli in a sound-attenuated booth. She neither hypo- nor hyper- articulated the tokens. Like the previous experiments, her speech was recorded through an AT4040 Cardioid Capacitor microphone with a pop filter in a sound-attenuated recording booth and amplified through an ART Tube MP microphone preamplifier (JVC RX 554V), digitized at 44K. She repeated all the stimuli ten times, and the five tokens of each form without any phonetic deviations (e.g. aberrant F0 contour, heavy creakiness, or clipping) were chosen.

#### 4.1.2 Noise and S/N-ratios

The noise used in this experiment was cocktail party noise, taken from the study used in Kawahara (2006). The reason for using this particular type of noise was to emulate the real communicative situations most closely. To obtain the cocktail party noise, Kawahara (2006) recorded a party using a SONY TCD-D8 portable DAT recorder. The recorded sound was divided into three-second noise stretches. Six such stretches were superimposed on top of one another.

Building on Binnie et al. (1974), the current experiment used three S/N-ratios: -6dB, -12dB, and -15dB where the signal dB was kept at the average of 60dB. Praat (Boersma & Weenink, 1999–2014) automatically adjusted the duration of the noise file to the duration of each stimulus

by the overlap-and-add method, and superimposed the adjusted noise file to each stimulus file.

#### 4.1.3 Procedure

Superlab (ver. 4.0, Cedrus) was used to present the stimuli. For each stimulus, possible responses given were binary. For example, for a sound stimulus [ $\alpha$ m], in one trial, the two visual responses were "am" or "an"; in the other trial, the two visual responses were "am" or "aŋ". This format allowed us to calculate the perceptual distance between any two minimal pairs differing in place. For each pair of visual cues, both possible orders were included in the test (e.g. "am" and "an"; "an" and "am"). The visual cue for [ $\eta$ ] was "ng".

The experiment started with a practice run in which the participants practiced the identification experiment, using a pair that differed in voicing, not in place. The practice session presented 10 items, and an experimenter stayed in the listening room so that the participants could ask questions after the practice run. The main session consisted of three blocks separated by a break sign. Each block contained all the stimuli for each S/N-ratio (9 target stimuli \* 5 tokens \* 2 visual cue combinations \* 2 visual cue orders = 180 tokens). All participants wore Sennheiser HD 280 Pro Headphones and used an RB-730 response button box (Cedrus) to register their responses. The order of the stimuli within each block was randomized by Superlab.

#### 4.1.4 Participants

Twenty-four native speakers of English participated in this study for course credits in linguistics or psychology classes. No participants who participated in the previous two experiments participated in this study. One speaker failed to respond to more than half of the trials, and hence this person's data was excluded.

#### 4.1.5 Analysis

We used a signal detection analysis to calculate the perceptual distance between each sound pair (Macmillan & Creelman, 2005). For each binary comparison, we calculated its d'-value, using z(Hit) - z(FalseAlarm). This signal detection analysis has an advantage of teasing apart sen-

sitivity, which reflects a perceptual distance, from bias, a listener's strategic bias to choose one option over the other (Macmillan & Creelman, 2005).<sup>3</sup> To analyze the d'-values statistically, a linear mixed model was run in which S/N-RATIO, MANNER, and PLACE were fixed factors.

#### 4.2 Results

Table 3 illustrates the average d'-values of each pair in Experiment III. The higher the d'-value, the more perceptible the pair was.

-6dB	Nasals	Voiced stops	Voiceless stops		
Labial vs. coronal	0.51 (0.37)	0.43 (0.37)	1.93 (0.27)		
Labial vs. dorsal	0.26 (0.25)	1.21 (0.38)	0.91 (0.42)		
Coronal vs. dorsal	0.34 (0.17)	1.25 (0.34)	2.37 (0.37)		
Averages	0.37	0.96	1.73		
-12dB	Nasals	Voiced stops	Voiceless stops		
Labial vs. coronal	0.11 (0.29)	0.24 (0.32)	2.02 (0.42)		
Labial vs. dorsal	0.21 (0.28)	0.76 (0.27)	0.93 (0.38)		
Coronal vs. dorsal	0.27 (0.39)	0.77 (0.31)	2.32 (0.36)		
Averages	0.20	0.59	1.76		
-15dB	Nasals	Voiced stops	Voiceless stops		
Labial vs. coronal	0.24 (0.25)	0.14 (0.31)	1.68 (0.44)		
Labial vs. dorsal	0.04 (0.20)	0.65 (0.36)	0.63 (0.34)		
Coronal vs. dorsal	-0.03 (0.36)	0.81 (0.31)	1.98 (0.49)		
Averages	0.09	0.53	1.43		

Table 3: The average d'-values in Experiment III (margins of errors for 95% confidence intervals).

A general linear mixed model analysis shows that all three factors had a significant impact on d'-values (S/N-RATIO: t = -4.11, p < .001; MANNER: t = 17.25, p < .001; PLACE: t = -2.35, p < .05). The lower the S/N-ratio (the noisier the noise), the lower the d'-values. PLACE showed its effect particularly in voiced stop pairs and voiceless stops pairs; in voiced stop pairs, labial-coronal pairs showed lower d'-values than the other two pairs; in voiceless stop

<sup>&</sup>lt;sup>3</sup>A question was raised whether A', a non-parametric measure of sensitivity, may have been a better measure. However, since Macmillan & Creelman (1996) show that A' is not free of distributional assumptions, we stick to more standard d'-measures.

pairs, labial-dorsal pairs showed lower d'-values than the other two pairs.<sup>4</sup> Most importantly in this context, MANNER is significant, showing the perceptibility hierarchy we observed in the previous two similarity judgment experiments: voiceless stops > voiced stops > nasals.

A contrast analysis comparing nasals and voiced stops shows that all three factors were significant (S/N-RATIO: t = -4.31, p < .001; MANNER: t = 6.42, p < .001; PLACE: t = 2.86, p < .01). Most importantly, nasals showed significantly lower d'-values than voiced stops. A contrast analysis comparing voiced and voiceless stops also showed that all three factors had a significant impact on d'-values (S/N-RATIO: t = -3.38, p < .001; MANNER: t = 10.97, p < .001; PLACE: t = -2.14, p < .001). As with the two previous similarity judgment experiments, the place contrasts are less salient in voiced stops than in voiceless stops.

#### 4.3 Discussion

To summarize, the identification experiment in noise shows the perceptibility hierarchy expected from the previous two experiments: voiceless stops > voiced stops > nasals, supporting the idea that nasals' place contrasts are weaker than oral stops' place contrasts. In fact, nasal place contrasts seem almost non-perceptible—i.e. d'-values are close to zero—under -12 dB and -15 dB S/N-ratio conditions. Indeed, the lower bounds of the 95% confidence intervals—the average values minus the margins of errors—overlap with zero in these conditions.

The current identification experiment thus yet again revealed a perceptibility difference of the place contrasts between nasal consonants and oral stops, supporting Jun's hypothesis (Jun, 1995, 2004). This result accords well with that of Pols (1983), but not with that of Winters (2002). The difference between the current experiment and that of Winter may have come from two sources. First, we used naturalistic sounds—both the targets and noise—to replicate the real communicative situations. In particular, noise was similar to those that speakers and listeners face in real communicative situations.

<sup>&</sup>lt;sup>4</sup>We do not have a good answer as to why voiced stops and voiceless stops show different patterns in terms of different place pairs. For voiceless pairs, labials and dorsals may have been most often confused because they are both grave consonants with concentration of energy in low frequency ranges (Jacobson et al., 1952). However, it is not clear why this confusion among two grave consonants does not extend to voiced stop pairs.

Second, the target consonants in the current experiment were placed in word-final position rather than in pre-consonantal position. The next experiment tested if the perceptibility differences observed in this experiment still hold in pre-consonantal position in which place assimilation occurs in phonology.

# 5 Experiment IV: Identification experiment in pre-consonantal position

#### 5.1 Introduction

The previous identification experiment shows that nasal place contrasts are less perceptible than oral consonant place contrats. The final question that we address is whether the same asymmetry holds in preconsonatal position, in which place assimilation actually occurs in phonology.

#### 5.2 Method

#### 5.2.1 Stimuli

To create preconsonantal environment, we first recorded the same speaker pronouncing [apə, atə, akə, abə, adə, agə] with stress on initial vowels. We then spliced off the initial stressed [a] vowels, and adjusted the amplitudes of the remaining portions—the unstressed second syllables—to 60dB.

We then concatenated each stimulus from Experiment III with the syllable that starts with a consonant that is non-homorganic to either of the two visual cues; for example, for the sound [am] whose two visual cues were "am" and "an", the concatenated CV syllable was [gə]; for the sound [ak] whose two visual cues were "at" and "ak", the concatenated CV syllable was [pə]. We chose non-homorganic consonants in order to prevent our listeners from defaulting to assimilated percept in the listening experiment (Beddor & Evans-Romaine, 1995; Kochetov & So, 2007; Malécot, 1956; Ohala, 1990).

Our pilot experiment shows that with a following CV syllable, the task is harder and in the

-15dB S/N-ratio condition, listeners would perform almost near chance in all three conditions. Therefore, we tested only -6dB S/N-ratio and -12 dB S/N-ratio condition. In this experiment, we repeated each token twice.

#### 5.2.2 Procedure

The procedure for this experiment is almost identical to that of Experiment III, except that the listeners were asked to identify the quality of initial syllables. The stimulus structure was as follows: for each S/N-ratio condition, we had 9 target stimuli \* 5 tokens \* 2 visual cues \* 2 orders \* 2 repetition=360 tokens. The order of the stimuli within each block was randomized by Superlab.

#### 5.2.3 Participants

Twenty-two students participated in this study for class credits in either linguistics or psychology classes. No participants who participated in the previous three experiments participated in this study.

#### 5.3 Results

Table 4 shows the average d'-values of each consonant pair in Experiment IV.

-6dB	Nasals	Voiced stops	Voiceless stops
Labial vs. coronal	0.29 (0.19)	0.12 (0.14)	0.98 (0.41)
Labial vs. dorsal	0.08 (0.20)	0.49 (0.32)	0.78 (0.35)
Coronal vs. dorsal	0.14 (0.17)	0.55 (0.24)	1.68 (0.54)
Averages	0.17	0.39	1.15
-12dB	Nasals	Voiced stops	Voiceless stops
Labial vs. coronal	0.00 (0.21)	0.00 (0.12)	0.73 (0.37)
Labial vs. dorsal	0.15 (0.15)	0.33 (0.23)	0.45 (0.20)
Coronal vs. dorsal	-0.11 (0.23)	0.07 (0.15)	1.63 (0.59)
Averages	0.01	0.13	0.93

Table 4: The average d'-values in Experiment IV (margins of errors for 95% confidence intervals).

A general linear mixed model analysis shows that S/N-RATIO and MANNER had a significant

impact on d'-values (S/N-RATIO: t = -3.05, p < .01; MANNER t = 11.28, p < .001; PLACE: t = 0.36, n.s.). The significant effect of S/N-RATIO shows that the louder the noise with respect to the targets, the lower the d' values. Most importantly, MANNER was significant because again we observe the following hierarchy: voiceless stops > voiced stops > nasals.

A contrast analysis comparing nasals and voiced stops shows that all the factors are significant (S/N-RATIO: t = -3.72, p < .001; MANNER: t = 2.98, p < .01; PLACE: t = 2.40, p < .05.). In this analysis, MANNER was significant (t = 2.98, p < .01), supporting the difference in the perceptibility of place contrasts between nasals and voiced consonants. The effect of PLACE is particularly visible in the -12dB condition in which the labial-dorsal pairs have high d'-values in both nasal pairs and voiced stop pairs.

A contrast analysis comparing voiced and voiceless stops shows that S/N-RATIO and MAN-NER had a significant impact on d'-values (S/N-RATIO: t = -2.61, p < .01; MANNER: t = 8.67, p < .001; PLACE: t = 0.53, n.s.). As in the other three previous experiments, the place contrast was better perceived in voiceless consonants than in voiced consonants.

#### 5.4 Discussion

We observe that the d'-values are generally lower in this experiment than in the previous experiment in which the target places were placed word-finally. This difference shows that the presence of a following consonant can mask the perception of coda consonants (even when the coda consonants' releases were not masked acoustically) (see Beddor & Evans-Romaine 1995 for a similar result). Most importantly, we again observe the perceptibility hierarchy: voiceless stops > voiced stops > nasals, except for one reversal in the labial-coronal pairs between nasals and voiced stops in -6dB SN ratio condition.

# 6 General discussion

#### 6.1 Summary

To summarize, all of the four experiments show the following perceptibility hierarchy of place contrasts: voiceless stops > voiced stops > nasals. The perceptibility differences were observed regardless of whether stops were clearly released (Experiment I, III, IV) or not (Experiment II). The differences were also observed in both clear listening environments (Experiments I and II) and in noisy environments (Experiments III and IV). The differences hold both in similarity rating experiments (Experiments I and II) and in identification experiments under noise (Experiments III and IV). Finally, the differences were observed both in word-final position (Experiments I-III) and in pre-consonantal position (Experiment IV).

The comparison between the two tasks—similarity judgment tasks and identification experiment in noise–also show that these two tasks reveal comparable results in terms of the perceptibility of contrasts (though see Babel & Johnson 2010) and moreover, that speakers can make conscious judgments about the perceptibility of contrasts (Steriade, 2008).

Overall, the current results are compatible with what is predicted by Jun's (1995, 2004) hypothesis that nasal place contrasts are perceptually weaker than oral stop place contrasts. More generally, the results are also compatible with the hypothesis that speakers are more willing to neutralize contrasts that are less perceptible (Boersma, 1998; Huang, 2001; Hura et al., 1992; Kawahara, 2006; Kohler, 1990; Lindblom et al., 1995; Steriade, 1997, 2001, 2008).

#### 6.2 **Remaining questions**

One remaining question is where the disagreement about the percetibility of place contrasts in the previous literature comes from, in particular the difference between the current results and Winters (2002). As discussed above, it could come from the difference in the kinds of noise that were used. The current experiment used naturalistic sounds—both the targets and noise—to replicate the real communicative situations. In particular, noise was similar to those that speakers and listeners face in real communicative situations. Therefore, we can conclude that Jun's hypothesis may be on

the right track, to the extent that speakers perceive nasal place contrasts less in a realistic speech setting. However, fully investigating the source of differences in the previous literature is beyond the scope of our paper.

Another question is why the nasal place contrasts are judged to be less distinct than the oral consonant place contrasts, and why the place contrasts were judged to be less distinct in voiced stops than in voiceless stops. For the first difference, Jun (1995; 2004) hypothesizes, following Malécot (1956), that coarticulatory nasalization in adjacent vowels blur the formant transition information, making the place contrasts in nasals less distinct. See also Fujimura (1962) for related observations about the acoustics of nasals, and Beddor & Evans-Romaine (1995) for more general discussion. Our experiment was not designed to test this hypothesis directly, and a future experiment is necessary.

For the second difference, it may be that since the pressure build-up behind the closure is stronger for voiceless consonants than for voiced consonants, bursts are stronger for voiceless consonants than for voiced consonants. Since bursts play an important role in cueing place distinctions (Kochetov & So, 2007; Malécot, 1956; Smits et al., 1996; Stevens & Blumstein, 1978; Tekieli & Cullinan, 1979; Winitz et al., 1972), stronger bursts of voiceless consonants may result in more distinct percepts. However, recall that in Experiment II, the difference in perceptual similarity still holds when we controlled for the amplitudes of releases. Alternatiely, Chen (1970) suggests that voiceless stops' closure is made with greater articulatory force and higher acceleration than voiced stops' closure, which may result in stronger formant transition cues. Admittedly, this hypothesis is speculative, and pursuing it further is beyond the scope of this paper.

A yet another limitation of this study is the fact that the participants of the current experiments are limited to the native speakers of English. There is of course thus a question of whether the current results hold for speakers of other languages. We hope that our experimental results are replicated with speakers of other languages.

#### 6.3 Phonetic perceptibility and phonological patterns

While the overall results support Jun's hypothesis at least under a noise that mimics realistic speech setting, we also find a perceptual asymmetry which is not necessarily reflected in phonological patterns: we consistently found that voiceless stop place contrasts are more salient than voiced stop place contrasts, but as far as we know, this difference is not reflected in phonology. It is possible that further investigation of place assimilation typology will reveal a language in which only nasals and voiced consonants assimilate, but to the extent that such a pattern is a true gap, then our results show that not all perceptibility differences can be reflected in phonology, i.e. that the perceptibility scales that underlie phonological patterns involve certain abstraction (Gordon, 2002; Kochetov, 2006; Kochetov & So, 2007). An important question to be addressed in future research is what distinguishes perceptibility differences that are reflected in phonology and those that are not.

One possible explanation may be that there is a feature—[son]—that distinguishes nasals on the one hand and voiced and voiceless obstruents on the other. On the other hand, there is arguably no feature that distinguishes voiceless stops on the one hand and voiced stops and nasals on the other, assuming that the [voice] feature for stops and the [voice] feature sonorants are different features, or that sonorants do not have [voice] feature at all (Mester & Itô, 1989; Rice, 1993). In this sense, there is no natural phonological class that distinguishes voiced stops and nasals, in exclusion of voiceless stops. This hypothesis is merely a speculation and needs to be tested in future research.

We admit that our experiments were not designed to address all of these questions, and the current paper indeed opens up many more research questions than it answers. However, it is not realistic to address all of these questions in one paper—we hope that more perception experiments will be conducted to address these issues. Nevertheless we hope to have offered one substantial step bearing on the issue of the perceptibility differences of place contrasts in nasals and oral stops, and its possible implication for phonological patterns of place assimilation. At the very least, the current experiments have shown that the prediction made by Jun (1995; 2004) can be confirmed in some experimental settings.

# References

- Baayen, Harald R. (2008) *Analyzing linguistic data: A practical introduction to statistics using R*.Cambridge: Cambridge University Press.
- Baayen, Harald R. (2009) LanguageR. R package.
- Baayen, Harald R., Doug.J. Davidson, & Douglas. M. Bates (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**: 390–412.
- Babel, Molly & Keith Johnson (2010) Accessing psycho-acoustic perception and language-specific perception with speech sounds. *Laboratory Phonology* **1**(1): 179–205.
- Bates, Douglas (2005) Fitting linear mixed models in R. R News 5: 27–30.
- Bates, Douglas, Martin Maechler, & Ben Bolker (2011) lme4: Linear mixed-effects models using S4 classes. R package.
- Beckman, Jill (1998) *Positional Faithfulness*. Doctoral dissertation, University of Massachusetts, Amherst.
- Beddor, Patrice & David Evans-Romaine (1995) Acoustic and perceptual factors in nasal place assimilation. *Rivista di Linguistica* **7**: 145–174.
- Benkí, José (2003) Analysis of English nonsense syllable recognition in noise. *Phonetica* 60: 129– 157.
- Binnie, Carl, Allen Montgomery, & Pamela Jackson (1974) Auditory and visual contributions to the perception of consonants. *Journal of Speech and Hearing Research* **17**: 619–630.
- Blevins, Juliette (2004) *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
- Blevins, Juliette (2006) A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* **32**(2): 117–166.
- Boersma, Paul (1998) Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives. The Hague: Holland Academic Graphics.
- Boersma, Paul (2008) Emergent ranking of faithfulness explains markedness and licesing by cue.Ms, University of Amsterdam.

Boersma, Paul & David Weenink (1999–2014) Praat: Doing phonetics by computer. Software.

- Byrd, Dani (1992) Perception of assimilation in consonant clusters: A gestural model. *Phonetica*49: 1–24.
- Chen, Matthew (1970) Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* **22**: 129–159.
- Cho, Young-mee Yu (1990) *Parameters of Consonantal Assimilation*. Doctoral dissertation, Stanford University.
- Ellis, Lucy & William Hardcastle (2002) Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics* **30**: 373– 396.
- Fleischhacker, Heidi (2001) Cluster dependent epenthesis asymmetries. In UCLA Working Papers in Linguistics 5, A. Albright & T. Cho, eds., Los Angeles: UCLA, 71–116.
- Fujimura, Osamu (1962) An analysis of nasal consonants. Journal of the Acoustical Society of America 34: 1865–1875.
- Fujimura, Osamu, M. J. Macchi, & L. A. Streeter (1978) Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. *Language and Speech* 21: 337–346.
- Gordon, Matthew (2002) A phonetically driven account of syllable weight. *Language* **78**(1): 51–80.
- Greenberg, Joseph & James Jenkins (1964) Studies in the psychological correlates of the sound system of American English. *Word* **20**: 157–177.
- Hayes, Bruce & Donca Steriade (2004) Introduction: The phonetic bases of phonological markedness. In *Phonetically Based Phonology.*, Bruce Hayes, Robert Kirchner, & Donca Steriade, eds., Cambridge: Cambridge University Press, 1–33.
- Holtman, Astrid (1996) *A Generative Theory of Rhyme: An Optimality Approach*. Doctoral dissertation, Utrecht Institute of Linguistics.
- Huang, Tsan (2001) The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. In *Ohio State University Working Papers in Linguistics 55: Studies on the Interplay of Speech Perception and Phonology*, Elizabeth Hume & Keith Johnson, eds., Columbus: OSU

Working Papers in Linguistics, 23–42.

- Huang, Tsan (2004) *Language-specificity in auditory perception of Chinese tones*. Ph.D. thesis, Ohio State University.
- Huang, Tsan & Keith Johnson (2010) Language specificity in speech perception: Perception of Mandarin tones by native and non-native speakers. *Phonetica* 67: 243–267.
- Hume, Elizabeth & Keith Johnson (2003) The impact of partial phonological contrast on speech perception. *Proceedings of ICPhS 2003* **XV**: 2385–2388.
- Hura, Susan, Björn Lindblom, & Randy Diehl (1992) On the role of perception in shaping phonological assimilation rules. *Language and Speech* **35**: 59–72.
- Jacobson, Roman, Gunner Fant, & Morris Halle (1952) *Preliminaries to Speech Analysis*. Cambridge: MIT Press.
- Jaeger, Florian T. (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* **59**: 434–446.
- Johnson, Keith (2003) Acoustic and Auditory Phonetics: 2nd Edition. Malden and Oxford: Blackwell.
- Jun, Jongho (1995) Perceptual and Articulatory Factors in Place Assimilation: An Optimality Theoretic Approach. Doctoral dissertation, University of California, Los Angeles.
- Jun, Jongho (2003) Positional faithfulness, sympathy and inferred input. Ms. Souel National University.
- Jun, Jongho (2004) Place assimilation. In *Phonetically based Phonology*, Bruce Hayes, Robert Kirchner, & Donca Steriade, eds., Cambridge: Cambridge University Press., 58–86.
- Kato, Hiroaki, Minoru Tsuzaki, & Yoshinori Sagisaka (1997) Acceptability for temporal modification of consecutive segments in isolated words. *Journal of the Acoustical Society of America* 101: 2311–2322.
- Kawahara, Shigeto (2006) A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* **82**(3): 536–574.
- Kawahara, Shigeto (2007) Half-rhymes in Japanese rap lyrics and knowledge of similarity. *Journal of East Asian Linguistics* **16**(2): 113–144.

- Kawahara, Shigeto (2009) Faithfulness, correspondence, and perceptual similarity: Hypotheses and experiments. *Onsei Kenkyu [Journal of the Phonetic Society of Japan]* **13**(2): 52–61.
- Kawahara, Shigeto & Kazuko Shinohara (2009) The role of psychoacoustic similarity in Japanese puns: A corpus study. *Journal of Linguistics* **45**(1): 111–138.
- Kochetov, Alexei (2006) Testing licensing by cue: A case of Russian palatalized coronals. *Phonetica* **63**(2-3): 113–148.
- Kochetov, Alexei & Connie K. So (2007) Place assimilation and phonetic grounding: A crosslinguistic study. *Phonology* **24**: 397–432.
- Kohler, Klaus (1990) Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In *Speech Production and Speech Modeling*, William J. Hardcastle & Alain Marchal, eds., Dordrecht: Kluwer, 69–92.
- Kurowski, Kathleen & Sheila E. Blumstein (1984) Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants. *Journal of the Acoustical Society of America* **76**: 383–90.
- Kurowski, Kathleen & Sheila E. Blumstein (1993) Acoustic properties for the perception of nasal consonants. In *Nasals, Nasalization and the Velum*, Marie Huffman & Rena Krakow, eds., New York: Academic Press, 197–224.
- Lindblom, Björn, Susan Guion, Susan Hura, Seung-Jae Moon, & Raquel Willerman (1995) Is sound change adaptive? *Rivista di Linguistitca* **7**: 5–37.
- Macmillan, Neil & Douglas Creelman (1996) Triangles in ROC space: History and theory of "non-parametric" measures of sensitivity and response bias. *Psychonomic Bulletin & Review* **3**: 164–170.
- Macmillan, Neil & Douglas Creelman (2005) *Detection Theory: A User's Guide. 2nd Edition*.Mahwah: Lawrence Erlbaum Associates Publishers.
- Malécot, André (1956) Acoustic cues for nasal consonants: An experimental study involving a tape-splicing technique. *Language* **32**: 274–84.
- Martin, Andrew & Sharon Peperkamp (2011) Speech perception and phonology. In *The Blackwell companion to phonology*, Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, & Keren Rice,

eds., Oxford: Blackwell-Wiley, 2334–2356.

- McCarthy, John J. (2011) Perceptually grounded faithfulness in Harmonic Serialism. *Linguistic Inquiry* **42**(1): 171–183.
- Mester, Armin & Junko Itô (1989) Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* **65**: 258–93.
- Mohanan, K. P. (1993) Fields of attraction in phonology. In *The Last Phonological Rule: Reflections on Constraints and Derivations*, John Goldsmith, ed., Chicago: University of Chicago Press, 61–116.
- Mohr, B. & W. S. Wang (1968) Perceptual distance and the specification of phonological features. *Phonetica* **18**: 31–45.
- Narayan, Chandra R. (2008) The acoustic-perceptual salience of nasal place contrasts. *Journal of Phonetics* **36**: 191–217.
- Nolan, Francis (1992) The descriptive role of segments: Evidence from assimilation. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, Gerard R. Docherty & Robert Ladd, eds., Cambridge: Cambridge University Press, 261–280.
- Ohala, John J. (1983a) The origin of sound patterns in vocal tract constraints. In *The Production of Speech*, Peter MacNeilage, ed., New York: Springer-Verlag, 189–216.
- Ohala, John J. (1990) The phonetics and phonology of aspects of assimilation. In *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, John Kingston & Mary Beckman, eds., Cambridge: Cambridge University Press, 258–275.
- Ohala, John J. & Manjari Ohala (1993) The phonetics of nasal phonology: Theorems and data. In *Nasals, Nasalization, and the Velum*, Marie Huffman & Rena Krakow, eds., New York: Academic Press, 225–249.
- Ohala, Manjari (1975) Nasals and nasalization in Hindi. In *Nasalfest*, Charles A. Ferguson, Larry M. Hyman, & John J. Ohala, eds., Language Universals Project, Stanford: Stanford University, 317–332.

Ohala, Manjari (1983b) Aspects of Hindi Phonology. Delhi: Motilal Banarsidass.

Padgett, Jaye (1995) Partial class behavior and nasal place assimilation. In Proceedings of the South

Western Optimality Theory Workshop 1995, Tuscon, AZ: The University of Arizona Coyote Papers, 145–183.

- Paradis, Carol & Jean-François Prunet, eds. (1991) The Special Status of Coronals: Internal and External Evidence. San Diego: Academic Press.
- Pols, Louis (1983) Three mode principle component analysis of confusion matrices, based on the identification of Dutch consonants, under various conditions of noise and reverberation. *Speech Communication* **2**: 275–293.
- R Development Core Team (1993–2014) *R: A language and environment for statistical computing*.R Foundation for Statistical Computing, Vienna, Austria.
- Repp, Bruno H. (1986) Perception of the [m]-[n] distinction in CV syllables. *Journal of the Acoustical Society of America* **79**(6): 1987–1736.
- Rice, Keren (1993) A reexamination of the feature [sonorant]: The status of sonorant obstruents. *Language* **69**: 308–344.
- Schane, Sanford (1972) Natural rules in phonology. In *Linguistic Change and Generative Theory*, Robert P. Stockwell & Ronald K. S. Macaulay, eds., Bloomington: Indiana University Press, 199–229.
- Singh, Sadanand & John Black (1966) Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *Journal of the Acoustical Society of America* **39**(2): 372–387.
- Smits, Roel, Louis Ten Bosch, & Rene Collier (1996) Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. *Journal of the Acoustical Society of America* 100: 3852–3864.
- Steriade, Donca (1997) Phonetics in phonology: The case of laryngeal neutralization. Ms. University of California, Los Angeles.
- Steriade, Donca (2001) Directional asymmetries in place assimilation: A perceptual account. In *The Role of Speech Perception in Phonology*, Elizabeth Hume & Keith Johnson, eds., New York: Academic Press, 219–250.

Steriade, Donca (2003) Knowledge of similarity and narrow lexical override. In Proceedings of the

29th annual meeting of the Berkeley Linguistics Society, Pawel M. Nowak, Corey Yoquelet, & David Mortensen, eds., Berkeley: BLS, 583–598.

- Steriade, Donca (2008) The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The nature of the word*, Kristin Hanson & Sharon Inkelas, eds., Cambridge [originally circulated in 2001 as ms. UCLA]: MIT Press, 151–179.
- Stevens, Kenneth & Sheila Blumstein (1978) Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* **64**: 1358–1368.
- Tekieli, Mary Ellen. & Walter L. Cullinan (1979) The perception of temporally segmented vowels and consonant-vowel syllables. *Journal of Speech and Hearing Research* : 103–121.
- Wang, W. S. Y. & C.J. Fillmore (1961) Intrinsic cues and consonant perception. *Journal of Speech and Hearing Research* 4: 130–136.
- Westbury, John R. (1979) Aspects of the Temporal Control of Voicing in Consonant Clusters in English. Doctoral dissertation, University of Texas, Austin.
- Westbury, John R. & Patricia Keating (1986) On the naturalness of stop consonant voicing. *Journal of Linguistics* **22**: 145–166.
- Winitz, Harris, M. E. Scheib, & James A. Reeds (1972) Identification of stops and vowels for the burst portion of /p,t,k/ isolated from conversation speech. *Journal of the Acoustical Society of America* **51**: 1309–1317.
- Winters, Stephen (2002) Perceptual influences on place assimilation: A case study. Ms. University of Calgary.
- Yu, Alan (2004) Infixing with a vengeance: Pingding Mandarin infixation. *Journal of East Asian Linguistics* 13: 39–58.
- Zwicky, Arnold (1976) This rock-and-roll has got to stop: Junior's head is hard as a rock. In *Proceedings of Chicago Linguistic Society 12*, S. Mufwene, C. Walker, & S. Steever, eds., Chicago: CLS, 676–697.
- Zwicky, Arnold & Elizabeth Zwicky (1986) Imperfect puns, markedness, and phonological similarity: With fronds like these, who needs anemones? *Folia Linguistica* **20**: 493–503.