DE GRUYTER MOUTON
DOI 10.1515/lp-2014-0003 — Laboratory Phonology 2014; 5(1): 37–68

Maria-Josep Solé

The perception of voice-initiating gestures

Abstract: This study examines how variation in production is perceived and then (re)interpreted by listeners, thus providing the link between phonetic variation and sound change. We investigate whether listeners can detect the nasal leak that may accompany utterance-initial voiced stops in Spanish, and reinterpret it as a nasal segment. Such reinterpretation may account for a number of sound patterns involving emergent nasals adjacent to voiced stops in oral contexts. Oral pressure, nasal/oral airflow, and audio were recorded for utterance-initial /b d p t/ produced by 10 Spanish speakers. Tokens showing different degrees of nasal leak (nasal C, maximum, medium, and no nasal leak) were placed intervocalically, where both /C/ and /NC/ may occur. The stimuli were presented to Spanish listeners for identification as /VNCV/ or /V(C)CV/. Identification results indicate a higher number of VNCV responses with incremental changes in nasal leak in voiced but not voiceless stimuli. Reaction time analysis showed shorter latencies to nasal identification for larger velum leak stimuli. The results suggest that listeners can 'hear' the nasal leak and fail to relate it to voicing initiation, interpreting a nasal segment. Thus a gesture aimed at facilitating voicing initiation may be interpreted as a new target goal.

Maria-Josep Solé: Universitat Autònoma de Barcelona. E-mail: mariajosep.sole@uab.cat

1 Introduction

While a large number of studies have looked at the variation in production which may give rise to sound change, fewer studies have looked at how this variation is perceived and (re)interpreted by listeners (but see, for example, Harrington et al. 2008; Recasens and Espinosa 2010; Beddor 2012; Harrington 2012). Moreover, most studies have focused on coarticulatory, aerodynamic, inertial, or rate-induced variation that, because it is automatic and predictable, could lead listeners to expect the resulting effects and correct for them (or fail to do so). This study focuses on the perception of variation resulting from implementational features, that is, features directed at achieving a certain goal, specifically, the use of nasal leak directed at achieving vocal fold vibration. Such implementational effects differ from regular phonetic variation in two important respects. First, they

are features *targeted* by the speaker (rather than mechanical); that is, they are programmed as part of the 'constellation of gestures' directed to achieve a certain acoustic effect (Browman and Goldstein 1989) – in this case, voicing during a stop. Second, they are *not fully predictable,* as they vary from speaker to speaker and segment to segment of the stop category, although they appear to be used rather consistently.

The present study reports a perceptual experiment designed to examine whether Spanish listeners can detect the nasal leak that may accompany utterance-initial voiced stops in Spanish and French (Solé 2011; Solé and Sprouse 2011), and reinterpret it as a nasal segment. Such reinterpretation would account for a number of cross-linguistic sound patterns involving the emergence of nasals adjacent to voiced, but not voiceless, stops in the absence of contextual nasals, e.g., prenasalized voiced stops, emergence of non-etymological nasals, voiced stops becoming nasals, or preservation of voicing exclusively after nasals (see Solé 2009 for a review).

The following section succinctly reviews articulatory and aerodynamic studies reporting the use of nasal leak to overcome the difficulty involved in achieving vocal fold vibration during a stop closure, especially in utterance-initial stops. Section 3 examines a number of phonological patterns illustrating the emergence of non-etymological nasals next to voiced stops; such nasals are hypothesized to emerge as a result of reinterpreting the nasal leak intended to facilitate voicing as a separate nasal segment. In Section 4 the results of a perceptual study that tests the sound change hypothesis directly are presented. Section 5 contains the discussion and conclusions.

2 Voicing during stops and nasal leak

The relative difficulty of maintaining voicing during a stop has been the object of extensive investigation (e.g., Rothenberg 1968; Ohala 1983; Westbury 1983; Westbury and Keating 1986) and is at the origin of the 'aerodynamic voicing constraint' (AVC). This constraint disfavors voicing in obstruents: the oral constriction associated with obstruents creates a build-up of intra-oral air pressure that reduces transglottal airflow and thus inhibits vocal fold vibration. Thus medial and final stops tend to devoice ('passive devoicing') after a few tens of milliseconds following the stop closure in the absence of additional articulatory adjustments.

In utterance-initial stops the aerodynamic and laryngeal conditions are less conducive to voicing than in medial or final stops, where voicing continues from the preceding vowel/sonorant. In utterance-initial stops, subglottal air pressure increases from atmospheric pressure in a characteristically linear manner, fol-

lowing a similar time course to the oral pressure increase during the stop constriction. Because oral and subglottal pressure rise in synchrony, the pressure differential is insufficient to set the vocal folds into vibration, and stop voicing is unlikely to occur without additional maneuvers. This difficulty is aggravated by the fact that after a pause the vocal folds must approximate and be duly tensed, and glottal vibration has to be initiated rather than sustained, which requires a greater pressure difference – 3–4 cm H_2O vs. 1–2 cm H_2O (Baer 1975: 139; Hirose and Niimi 1987) – due to the need to overcome inertial effects. As a result, closure voicing during an utterance-initial stop is less likely to occur and is typically shorter than after a vowel (Westbury and Keating 1986).

Previous studies have provided evidence of a number of ways in which speakers may reduce the build-up of oral pressure during stop closures, and thus allow phonation to initiate earlier and/or last longer. Three basic mechanisms have been reported (Rothenberg 1968; Westbury 1983):

- passive expansion of the oropharyngeal walls by reducing the level of muscle contraction
- active enlargement of the oral cavity (lowering the larynx, elevating the soft palate, advancing the tongue root, depressing the tongue body)
- releasing airflow through an incomplete oral or velopharyngeal closure

The specific adjustments used vary within and across speakers and segments, and possibly across languages (Bell-Berti 1975; Westbury 1983; Solé 2011).

In this paper we focus on (nasal) airflow being released from the oral cavity through an incomplete (or delayed) velum closure during voiced stops. Venting air through the nose reduces oral pressure, hence facilitating transglottal flow and voicing during a stop. Using a variety of techniques, evidence for nasal leak (and thus open velum) during voiced but not voiceless stops has been found in German (Muselhold 1913; Pape et al. 2006, acoustic data), American English (Yanagihara and Hyde 1966, nasal flow data; Rothenberg 1968; Kent and Moll 1969, cinefluorography), Hindi and Telugu breathy voiced stops (Rothenberg 1968), and Sindhi voiced stops (Nihalani 1975). (See further references in Ohala and Ohala 1991).[1] In addition, greater velopharyngeal closure

1 While the commonly accepted view is that nasal leak is an adjustment used by speakers to initiate or prolong vocal fold vibration during a stop closure, Lubker (1973) and Westbury (1983) suggest that the nasal flow differences between voiced and voiceless stops are small enough to be attributed to velum elevation – that is, the pumping action of the closed and rising velum – rather than velopharyngeal opening. Cineradiographic data by Kent and Moll (1969) showing differences in velopharyngeal closure in voiced and voiceless stops lend support to the traditional view.

force has been found for voiceless than voiced stops (Kuehn and Moon 1998), suggesting that the velic valve is finely controlled as a function of stop voicing.

A series of studies by Ohala and colleagues (Ohala and Sprouse 2003; Solé et al. 2008) introduced changes in oral pressure during speech with a pseudo-nasal valve and evaluated the effects on stop voicing. Two methods were used to simulate nasal leak, a 'buccal vent' (a large diameter catheter inserted via the buccal sulcus into the pharyngeal region, behind the stop constriction), and a 'nasal vacuum' (a small catheter inserted into the pharynx via the nasal passage and connected to a vacuum source that actively sucks air out). In both methods the external end of the catheter was randomly opened or closed by the experimenter. Ohala and Sprouse (2003) found that devoiced velar stops became voiced when oral pressure was reduced with the nasal vent. Solé et al. (2008) examined the effects of pressure variations using a 'buccal vent' on Voice Onset Time of stops in utterance-initial position (e.g., b̲oy), final position (e.g., la̲b), and in stop clusters (e.g., la̲b b̲oy). They found that when pressure is reduced due to nasal venting, voicing (1) begins earlier in utterance-initial stops (pressure values below 3–4 cm H_2O result in prevoiced stops) and (2) is prolonged in final stops and voiced stop clusters. Their results show that American English voiced stops were passively devoiced: speakers had an adducted (closed) glottis during the production of stops but lacked a sufficient transglottal pressure differential. When oral pressure was reduced due to nasal leak, and thus the pressure differential for transglottal flow increased, vocal fold vibration resulted.

Using Sprouse's (2008) aerodynamic model, Solé et al. (2008) also simulated nasal leak by varying the magnitude and timing of velic closure during voiced stops and derived the variations in air pressure and flow and, as a consequence, in voicing.[2] For utterance-initial stops, if nasal leak was present during the first part of the stop closure (the velum closed near the end of the stop to allow pressure to build up for the release), the oral pressure increased at a slower rate than it would without nasal leak, and thus the pressure difference was sufficient to initiate voicing during the closure (i.e., negative VOT). In a similar way, for final stops, nasal leak during the closure prevented oral pressure from rising as rapidly as it would in the absence of a leak, and voicing was maintained throughout the closure.

More recently, Solé (2011) and Solé and Sprouse (2011) examined cross-language differences in the voicing of utterance-initial voiced stops and in the use

2 Even though the model does not yet include vocal fold vibration, it can indicate when the critical level for voicing is reached.

of active maneuvers – such as nasal leak, oral leak, and oral cavity enlargement gestures – to achieve closure voicing. They collected aerodynamic and acoustic data for utterance-initial /b d p t m/ for speakers of Spanish and French, with typically prevoiced stops, and speakers of English, with typically devoiced stops. They found that the majority (>85%) of Spanish and French voiced stops showed voicing lead, and involved some type of adjustment, with nasal leak (as determined by the occurrence of nasal flow during the stop closure) being the most common (54% and 62% of the prevoiced tokens in Spanish and French, respectively). In contrast, in English the great majority (83.2%) of initial voiced stops were devoiced, and the aerodynamic data reflected fewer cases of voice facilitating adjustments than in the Romance languages. Furthermore, quantitative analysis for the Spanish data showed a correlation between nasal leak (or delayed velic closure) and voicing during the stop closure. First, the time between oral and nasal closure was longer for phonologically voiced than voiceless stops and, among the former, for prevoiced than devoiced stops. Second, velopharyngeal closure was related to phonation onset such that as voicing was initiated, the velum began to raise. These results suggest that occurrence of velum leakage is related to vocal fold vibration.

3 Sound patterns

Sound patterns involving the emergence of nasal consonants adjacent to voiced stops are rather common. A number of patterns involving emergent nasals next to voiced, but not voiceless, stops in the absence of a contextual nasal are illustrated below (see Solé [2009] for an extensive review). For descriptive purposes, we will group the various cases into three main patterns: (1) prenasalized voiced stops, (2) the emergence of non-etymological nasals, and (3) the substitution of voiced stops by nasals. In addition, the relationship between the distribution of emergent nasalization and the difficulty in initiating/maintaining vocal fold vibration in certain contexts will be illustrated. Note that in all these cases, nasals emerge adjacent to voiced stops in the absence of contextual or etymological nasals. These patterns, therefore, cannot be accounted for by acoustic-auditory factors (e.g., the lower tolerance of voiceless vis-à-vis voiced stops to nasalization, due to the high pressure required for a noisy release burst for voiceless stops, along the lines of Ohala and Ohala 1991, 1993), the *NÇ constraint (Pater 1999), or post-nasal voicing (Hayes and Stivers 2000), simply because these explanations rely on the detrimental effect of nasals on adjacent (mostly voiceless) obstruents. The patterns presented here, however, do not

contain an etymological nasal or occur in a nasal context, and therefore a different explanation must be sought.

– *Prenasalized voiced stops*. Languages with contrastive prenasalized voiced stops and voiceless stops (but no simple voiced stops), e.g., languages in Austronesia, Papua, and South America (Maddieson and Ladefoged 1993: 256), are illustrated in (1). Such prenasalized stops have been analyzed as the voiced stop series (for example, in Mixtec; Piggot 1992; Iverson and Salmons 1996).

(1) Waris (Brown 2001)

[ᵐb] *b̲anda* 'snake', *tom̲b̲ol* 'dry'	[p] *p̲anda* 'pitpit type', *nop̲o* 'eye'
[ⁿd] *d̲amba* 'tree sp.', *wan̲d̲* 'pitpit grass'	[t] *t̲at̲a* 'meat', *lot̲* 'banana type'
[ᵑg] *g̲o* 'go', *en̲g̲ala* 'hand'	[k] *k̲ao* 'tree sp.', *ok̲ala* 'distant'

The UPSID database contains 19 languages with prenasalized segments, all of which are voiced obstruents, overwhelmingly stops (Maddieson 1984: 67).

– *Phonetic pre- and post-nasalization of voiced but not voiceless stops in an oral context*. Optional prenasalization of voiced, but not voiceless, stops occurs in Bola (Wiebe 1997), Tok Pisin (Smith 2002), and Awara (Quigley 2003: 16), exemplified in (2).

(2) Awara[3]

/b̲am/ [bam] ~ [ᵐbam] 'log'	/teb̲ana/ [tɛmbana] 'morning'
/d̲ajip/ [dajip] ~ [ⁿdajip] 'look at them'	/gad̲oŋ/ [gandoŋ] ~ [ᵑgandoŋ] 'grass shoot'
/gʌlʌŋ/ [gʌlʌŋ] ~ [ᵑgʌlʌŋ] 'hook'	/jag̲ʌ/ [jaŋgʌ] 'water'

Postnasalization of final voiced (but not voiceless) stops has been reported in Lancashire dialects of English, illustrated in (3), in the South American languages Kaingang and Içua Tupí (Herbert 1986: 206, cited in Laver 1994: 232), in Vietnamese dialects, and in Senegalese Wolof (Ward 1939, cited in Jones 2001). The emergence of a nasal that is homorganic with the stop reflects that the velum is lowered during the latter part of the stop closure, resulting in a nasal release.

3 In Awara intervocalic voiced stops always follow a homorganic nasal while voiceless stops do not.

(3) Lancashire dialects (Jones 2001)
[ˈspɪt ə ˈgɒbm] ‘spit a gob (phlegm)’
[ˈkɔːv ə ðɪ ˈlɛgŋ] ‘calf of thy leg’
[ˌuːz ˈwɛdn] ‘she’s wed (married)’

– *Nasalization of voiced but not voiceless stops in an oral context*. Voiced but not voiceless stops become nasals in an oral context, e.g., Old Irish (Hickey 1997) and the Po-ai dialect of Tai, illustrated in (4).

(4) Dialects of Tai (Li 1977: 68–69, 107):

Proto-Tai	Siamese	Lungchow	Po-ai	
**ʔb-*	*baa*	*baa*	*maa*	‘shoulder’
**ʔb-*	–	*buu*	*muu*	‘onion’
**ʔd-*	*dïat*	*dïït*	*naat*	‘hot’

These cases neither contain a nasal etymologically nor occur in a nasal context, and the pre-/post-nasalized stop can therefore not simply be the result of ‘preservation’ of historical traces or coarticulatory nasalization. The spontaneous nasalization of voiced stops illustrated in these patterns most likely reflects the phonologization of the nasal leak used to facilitate voicing in stops.

Consistent with this hypothesis, prenasalization (examples (1) and (2)) is found word initially, and therefore utterance initially, where phonation is more difficult to initiate – due to the lowered subglottal pressure utterance initially and the larger pressure difference required to set the vocal folds into vibration compared to that required to sustain voicing. In contrast, postnasalization (example 3) occurs word and phrase finally, where voicing is difficult to sustain due to oral pressure build-up over time and decreasing subglottal pressure (Westbury and Keating 1986; Slifka 2000).

Furthermore, the hypothesis that nasal leakage is a maneuver to facilitate voicing in stops is supported by the distribution of spontaneous nasalization, that is, the tendency for nasals to emerge (or to be preserved longer) in contexts where vocal fold vibration is more difficult to maintain/initiate:

– in velars (with a smaller back cavity volume and lesser area of compliant tissue) vis-à-vis stops at more anterior places (Ohala 1983)
– in phrase-initial vis-à-vis medial position (Westbury and Keating 1986)

Evidence for the former comes from Japanese, where prenasalization of voiced stops was first lost in labials in Middle Japanese, later in coronals, and only recently in velars (Yamane-Tanaka 2005). The Mixtec data in (5) show a

similar distribution of prenasalization in voiced stops: velars and alveolars are prenasalized, but labials are not (with only optional prenasalization word initially in Chalcatongo Mixtec). Evidence for the latter (i.e., phrase position effects) also comes from Mixtec dialects in (5): more prenasalization is found word and phrase initially than word medially, at least in the bilabial series. Voiced velars only occur when the aerodynamic conditions conducive to voicing are maximized: word medially and prenasalized.

(5) Distribution of prenasalization in word-initial and medial position in Mixtec varieties (Iverson and Salmons 1996)

Word-initial	Word-medial
$^{(m)}$b	b
^{n}d	nd
-	ŋg

4 Perceptual study

We have shown that the effects of voicing implementation using incomplete nasal closure are consistent with the phonological patterns of emerging nasals. However, the voicing implementation account 'explains' the sound patterns involving nasals only if the listener can detect the nasal leak and reinterpret it as a nasal segment, that is, if the speaker fails to associate the consequences of the nasal gesture with a voiced stop and instead attributes the spectral consequences of nasalization to an intended nasal segment. A perceptual experiment was designed to obtain empirical evidence of such misperceptions in Spanish listeners. Spanish was chosen because, as with other Romance languages, it uses voicing during the closure to cue the voiced-voiceless contrast. Spanish has been shown to use nasal leak rather consistently to achieve closure voicing (Solé and Sprouse 2011), and, consequently, listeners must be familiar with the occurrence of this adjustment in voiced stops.

A perception test with natural speech stimuli was designed. Test syllables containing utterance-initial voiced stops with different degrees of velum leak were excised from the original sentences, placed in intervocalic position (where both -C- and -NC- can occur), and presented to listeners for identification. The hypothesis to be tested was that, if listeners hear the nasal leak but fail to attribute it to voicing initiation, they would tend to identify a nasal consonant. It was also predicted that they would tend to report more nasals with incremental velum leak.

4.1 Method

4.1.1 Test materials

Oral pressure, nasal airflow, oral flow, and audio were recorded for utterance-initial /b/, /d/, /p/, /t/, and /m/ produced by 10 Spanish speakers – 3 female (F1, F7, F9) and 4 male speakers (M3, M4, M5, M6) of continental Spanish; 1 female Mexican speaker (F2); and 1 female and 1 male Uruguayan speaker (F8, M10) – with each token produced 10 to 13 times. This was a subset of the data analyzed for production in Solé (2011) and Solé and Sprouse (2011) (see those works for further experimental details). Velar consonants were not included because the tube measuring oral pressure interfered with the production of back consonants. The subjects were instructed to read the following sentences as if they were a dialogue between A and B in order to obtain two isolated utterances, with the segment of interest being the beginning of the second utterance (the /b/ in Bárbara, the /d/ in Débora, etc.). The two sentences (A and B) were presented sequentially on a computer screen, such that B was presented only after A had been produced, to ensure a pause between the two sentences.

A: ¿Cómo se llama ella? ('What's she called?')
B: B̲árbara [D̲ébora, P̲aula, T̲ábata, M̲arta].

All sentences were sampled at 22.050 Hz. A first qualitative analysis of the data showed two main patterns of prenasalization in voiced stops. The first pattern involved *delayed* velic closure relative to the oral closure, that is, velopharyngeal closure following oral closure. This pattern is illustrated in Figures 1a and 1b. Oral pressure build-up for stops typically begins when the oral and nasal valves are closed. Delayed closing of the nasal valve results in a nasal leak during the closure that slows down oral pressure build-up and helps achieve the transglottal pressure difference needed for voicing. Once voicing is initiated, the velum closes, pressure rises rapidly, and the amplitude of voicing diminishes. The waveform illustrates the increasing amplitude of glottal pulsing during the prenasalized portion, reflecting the increased flow through the glottis due to nasal leakage (as opposed to decreasing amplitude of voicing in the latter part of the stop with the nasal valve closed). During the first portion of the stop, these tokens typically exhibit a slow pressure rise (or lack thereof as in Figure 1) and concurrent nasal flow. The magnitude of nasal leakage (i.e., volume of air flowing out of the nose) during the initial part of the stop closure tends to be comparable to the volume of nasal flow at the end of the utterance (see Figures 1a and 1b) and to the /m/ in 'llama' in the carrier phrase.

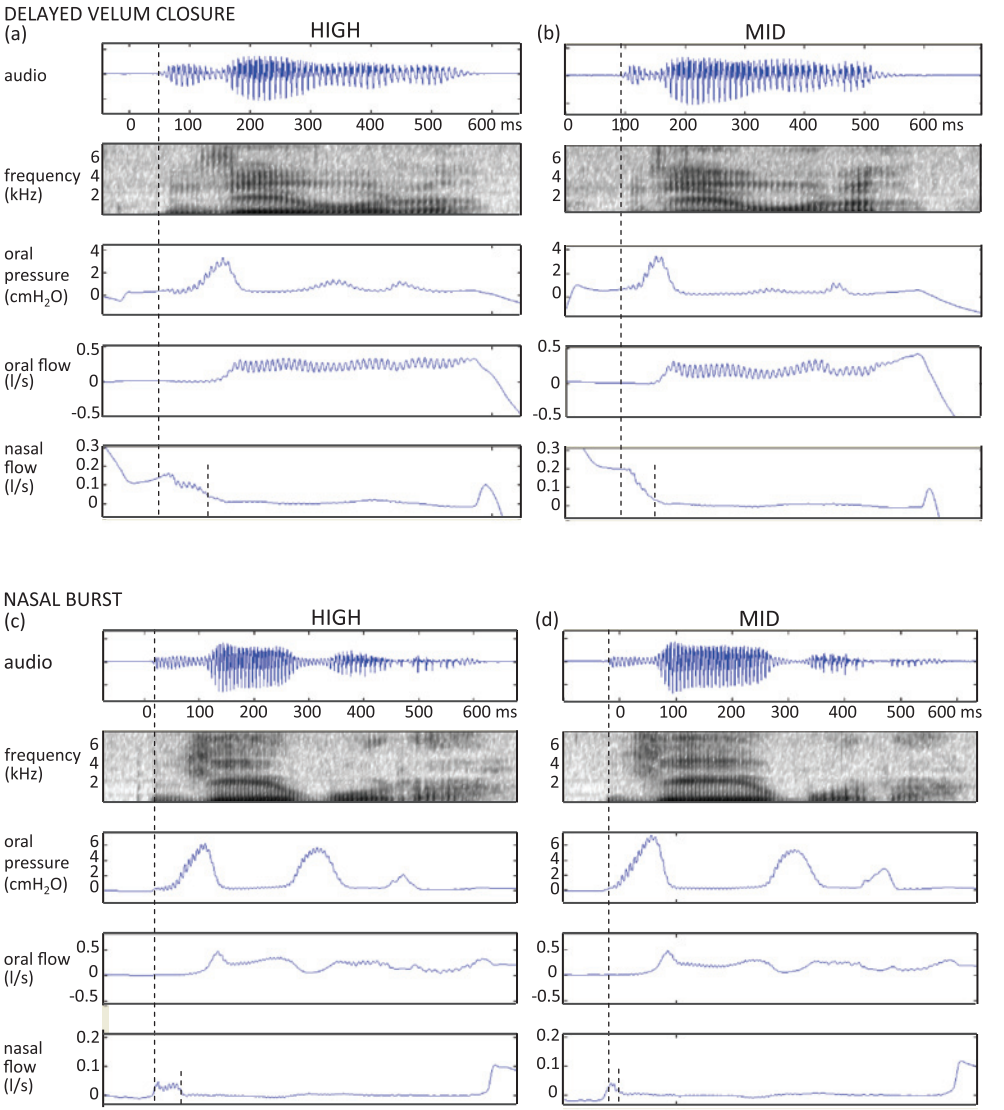

Fig. 1: Waveform, 0–7 kHz spectrogram, oral pressure (in cmH_2O), oral flow, and nasal flow (in l/s) for high and mid 'delayed velum closure' tokens (a and b), and high and mid 'nasal burst' tokens (c and d) of Spanish *Débora* [ˈdˠeβ̪ora]. The long vertical line indicates onset of vocal fold vibration (oral closure in place). The short vertical line (nasal flow to zero) indicates complete velum closure.

In a second pattern, illustrated in Figures 1c and 1d, the oral and nasal valves both close (i.e., the oral and nasal flows drop to zero), but a subsequent brief opening of the velum (seen as a *burst* of nasal airflow) accompanies the initiation of voicing. After voicing initiates, nasal flow drops to zero. That is, a momentary nasal leakage helps kick-start voicing. Voiceless stops do not show such nasal burst. Both types of prenasalization diminish oral pressure and facilitate the initiation of voicing. As soon as voicing is initiated, the velum typically closes. While speakers show a preference for using one type or the other, and some subjects use one type exclusively, half use both. There appears to be no difference in the frequency of occurrence of prevoicing in the two patterns, and they do not appear to be associated with place of articulation of the stop (labial vs apical).

In order to test the perceptual consequences of the two patterns found in production, two incremental nasalization continua were constructed with natural speech stimuli: a 'delayed velic closure' and a 'nasal burst' continuum. It was hypothesized that listeners would be more likely to detect a nasal segment in tokens showing maximum delayed velic closure/largest nasal burst than in tokens with less velum leak. A second hypothesis was that, for the same magnitude of nasal leak, listeners would be more likely to report a nasal before voiced than voiceless stops. This is because nasal leak has no acoustic consequences before voiceless stops (except possibly a less intense burst); that is, without a glottal source that resonates partly in the nasal cavity, nasal leak is not acoustically detectable.[4]

The perception stimuli were selected as follows for the delayed velic closure continuum. Measurements were made of duration of stop voicing, peak oral pressure, and delayed nasal closure (defined as the time interval between oral closure[5] and velum closure, indicated by nasal flow to zero) in initial stops in order to identify three tokens each with maximum, medium, and no delayed velic closure (i.e., no nasal leak) for each segment (/b d p t/). (See Figure 1 for an example of high (a) and medium (b) leak tokens.) These tokens were selected for the high, medium, and no delayed velic closure condition. Because perceptual results are

4 One reviewer raised the question of whether something like a voiceless nasal could be heard before voiceless stops. The answer is 'no'. Recall that the source of sound for voiceless nasals is frication at the nostrils; it appears that nasal leak was not sufficiently large to create noticeable turbulence at the nostrils.

5 The criteria used to determine the beginning of the oral closure were the occurrence of one of the following indicators: (1) oral airflow to zero and pressure build-up. If oral airflow dropped to zero (i.e., the oral articulators closed) at the end of the first phrase without a rise in pressure, then (2) beginning of oral pressure build-up, or (3) onset of glottal vibration (in Spanish voicing may begin without a rise in pressure, as illustrated in Figure 1).

highly dependent on specific phonetic details of the stimuli, three tokens were chosen for each step change. High velum leak was taken to occur when delayed nasal closure exceeded 50 ms (but the velum closed before stop release), medium when it was 20–50 ms, and no velum leak when nasal closure was close to or preceded oral closure (around 0 ms, but generally negative values); see Table 2. It is worth noting that although the range of values of oral closure-nasal closure for the three conditions was the same for voiced /b d/ and voiceless /p t/, the mean values were somewhat larger for voiced than for voiceless stops (in particular, for the high nasal leak condition). This was because voiceless tokens clustered around the low end of the range, whereas voiced tokens clustered around the high end. This trend is in agreement with reports in the literature (e.g., Ohala and Ohala 1991) indicating that in N-stop sequences velic closure is completed earlier before voiceless than voiced stops (presumably due to the lower tolerance of voiceless stops to velum leakage, which bleeds the oral pressure required for a noisy release burst). We chose to include the naturally observed values (e.g., maximum nasal leak values for voiced and voiceless stops rather than intermediate values for voiced stops and maximum values for voiceless stops) in order to create a large effect that allowed listeners to detect velum leak, and to simulate naturally occurring conditions.

Tokens showing a 'nasal burst' accompanying voicing initiation were also selected for the perceptual test. Only voiced stops showed such nasal bursts, and, therefore, only /b/ and /d/ tokens were selected for the incremental 'nasal burst' condition. Two tokens each showing a nasal burst with long duration,[6] medium duration, and with no nasal burst for each segment were selected for the high, medium, and no nasal burst conditions. An illustration of the high and medium nasal burst tokens is shown in Figures 1c and 1d. The duration of the nasal flow exceeded 40 ms for the high nasal burst condition and ranged between 20 and 40 ms for the medium burst condition. A summary of the experimental conditions is presented in Table 1.

The perception test included three tokens each for the high and medium delayed nasal closure conditions for voiced segments and two tokens for each of the other conditions. They were sampled from productions by speakers F1 (7), F2 (3), M4 (2), M5 (2), M6 (8), F7 (2), F8 (2), F9 (3), and M10 (7). Whenever possible tokens from the same speaker were selected for the different levels within a condition (e.g., high, medium, no delayed nasal closure). However, because speakers

6 An attempt was made to select tokens by magnitude (rather than duration) of nasal flow, but it was found that the magnitude of the nasal burst did not vary significantly for different durations, while it varied considerably from speaker to speaker.

Table 1: Nasal lags and duration of segments in the /NC/ (nasal + voiced/voiceless stop) for the experimental conditions. Values are in ms.

Condition	Delayed nasal closure		Nasal burst
	/b/ /d/	/p/ /t/	/b/ /d/
No	≤0		≤0
Medium	20–50		20–40
High	>50		>40
NC	70+50	55+65	70+50

tended to use the same nasal-timing pattern (e.g., delayed nasal closure) for all tokens, tokens from different speakers were used for the 'no' conditions. Table 2 presents the aerodynamic and acoustic characteristics of the 36 tokens chosen for perceptual evaluation. The peak oral pressure value for each token is also reported. Indeed, prenasalization is usually associated with a lower pressure build-up compared to tokens without velum leak, because air is escaping through the velopharyngeal port; however, other factors may also influence pressure values.

To create the perceptual stimuli, the initial /C_{stop}V(C)/ portion of [Bár]b*a*ra, [Dé]b*o*ra, [Pau]l*a*, [Tá]b*a*ta was excised from the original sentences and a vowel was added at the beginning of the sequence, such that the perceptual stimuli had a /VC_{stop}V(C)/ structure. The vowel was excised from the second syllable of the same token (shown in italics above) and consisted of that portion of the waveform in which none of the consonantal information could be perceived (though some VC transitions must have remained, as addressed below). The initial vowel was included to place the segment of interest in intervocalic position where both /C_{stop}/ and /NC_{stop}/ may occur.

The initial vowel was added at a fixed time (120 ms) before stop release such that the intervocalic portion which was to be identified as /mb/ or /b/ had a constant duration. This procedure was preferred to the alternative method of adding the initial vowel at a fixed time before phonation onset because duration of voicing varied from token to token (with tokens with more prenasalization tending to have longer voicing duration), and this would have resulted in different durations of the intervocalic portion (longer for tokens with longer voicing). If so, listeners could be identifying /mb/ or /b/ based on durational differences rather than nasal leak. In addition, adding the vowel at a constant time before stop release, rather than before voicing onset, allowed us to keep the time between vowels constant for voiced and voiceless tokens

The decision to add the initial vowel 120 ms before stop release was based on the following considerations: onset of voicing for phrase-initial /b d/ typically

Table 2: Duration of nasal flow (i.e, oral-to-nasal closure), peak oral pressure, and duration of voicing for the tokens selected as perception stimuli in the delayed nasal closure and nasal burst conditions. Each token was presented 2 or 3 times as specified. The no delayed nasal closure tokens were also used for the no nasal burst condition.

Segment	Speaker	Condition	Duration velum leak (ms)	Oral pressure (cmH_2O)	Duration of voicing (ms)	Number of repetitions	Label
			Delayed nasal closure				
/b/	F1	high	65.8	1.27	80.8	3	bH
	M6	high	82.1	3.16	116	3	bH
	M10	high	51	2.86	79.8	3	bH
	M6	mid	40.2	3.13	70.7	3	bM
	F9	mid	48.9	3.06	106	3	bM
	F9	mid	46.2	4.38	85.4	3	bM
	F8	no	0	3.89	65.2	2	b-no
	M10	no	−30	2.07	68.8	2	b-no
/d/	F1	high	62	2.01	82.75	3	dH
	M6	high	63.8	3.2	99.2	3	dH
	M10	high	50.6	6.58	120.3	3	dH
	F1	mid	33	2.9	70.8	3	dM
	M6	mid	37.75	4.3	66	3	dM
	F9	mid	33.5	5.54	94	3	dM
	M2	no	−30	7.8	139	2	d-no
	F8	no	−8	2.16	53.8	2	d-no
/p/	F1	high	52.5	3.89		2	pH
	M6	high	52	5.26		2	pH
	F1	mid	31.8	4.6		2	pM
	M10	mid	30	6.28		2	pM
	M6	no	0	6.16		2	p-no
	M10	no	−33	5.3		2	p-no
/t/	F1	high	58.7	5.7		2	tH
	M10	high	54	6.64		2	tH
	M6	mid	26	6.98		2	tM
	M6	mid	38	6.6		2	tM
	F1	no	−8	5.58		2	t-no
	M10	no	12	7.11		2	t-no
			Nasal burst				
/b/	F2	high	51.4	5.3	67.8	3	$b_{burst}H$
	M4	high	42	7.5	77.6	3	$b_{burst}H$
	F2	mid	24.25	5.87	45	3	$b_{burst}M$
	M4	mid	22.2	5.7	56.2	3	$b_{burst}M$

Table 2 (cont.)

Segment	Speaker	Condition	Duration velum leak (ms)	Oral pressure (cmH_2O)	Duration of voicing (ms)	Number of repetitions	Label
/d/	M5	high	45	3.53	100	3	$d_{burst}H$
	F7	high	40	4	73	3	$d_{burst}H$
	M5	mid	24.6	2.16	68.7	3	$d_{burst}M$
	F7	mid	31.25	4.33	58.25	3	$d_{burst}M$

occurred 60 ms before stop release in Spanish;[7] if a preceding nasal consonant were present, it would have a considerably long duration, approximately 100 ms (the ratio N-to-D in ND clusters is approximately 2:1 for French [Cohn 1990] and 1.5:1 for Catalan [Llach 1999; Recasens et al. 2004]). Therefore, adding the vowel 60 ms before the voiced stop (Gap-to-D ratio 1:1) was considered to be an ambiguous value that could allow, but not cue, the identification of a nasal. A 120 ms gap was also considered adequate for voiceless tokens (/p t/) (ratio of N-to-T in NT clusters is 0.8:1 [Llach 1999; Recasens et al. 2004]).

One potential problem is that when utterance-initial voiced stops (supposedly realized as full-closure stops) are placed in intervocalic position when a vowel is added, such intervocalic stops would be spirantized [β ð ɣ] in Spanish. However, our data show that, in accordance with reported gradient behavior in the stop-approximant allophonic alternation in Spanish (Torreblanca 1979: 461; Cole et al. 1999; Soler and Romero 1999; Ortega-Llebaria 2004), approximately half of the utterance-initial stops were spirantized, mostly /d/ tokens[8] (see Solé 2011). Furthermore, a pilot experiment showed that the resulting /V+C_{stop}V/ sequences – most likely due to residual VC transitions in the initial vowel – were sometimes heard as /VC+CV/, e.g., 'agba', 'adba' (which are licit stop clusters in Spanish, involving either a full closure or an incomplete closure, e.g., *advertir* [db], [ðβ] 'to warn', *abdicar* [bd], [βð] 'to abdicate', *magdalena* [gd], [ɣð] 'muffin'). Thus, listeners were asked to identify the stimuli as /VNCV/ or /V(C)CV/. For these reasons we feel confident that the nonsense sequences presented to listeners were natural in Spanish.

7 This was a longer duration than the reported values for voiced stops in medial ND clusters for comparable languages (Cohn 1990: 102, 107–108 for French; Recasens et al. 2004 and Llach 1999 for Majorcan Catalan).

8 Spirantization of initial /d/ is illustrated in Figure 1. Oral airflow and high-frequency noise can be observed during the latter part of the stop constriction.

In order to provide listeners with a reference token of the /NC/ identification response (and not to bias them against this response), one token each of the sequences /VmbV/, /VndV/, VmpV/, /VntV/ was added. The NC stimuli were created by adding a portion of nasal murmur from the carrier sentence before the stop in the ratios observed for ND and NT clusters in a related Romance language, Catalan: for /VmbV/, /VndV/ stimuli, 70 ms nasal murmur + 50 ms voiced stop closure (1.4 ratio); /VmpV/, /VntV/: 55 ms nasal murmur + 65 ms voiceless stop closure (.85 ratio) (Llach 1999; Recasens et al. 2004 for Majorcan Catalan).

One concern related to perceptual testing based on splicing word onsets into a medial context is that nasal leakage seems to be rarer word medially because ongoing voicing requires less support than initiation of voicing. Thus, Spanish listeners may not have developed a perceptual parsing strategy assigning nasal leakage to voicing in this context, and, consequently, there may be a bias towards perception of a nasal segment. We think that this potential bias is minimized because nasal leakage has been reported when a voiced stop follows a vowel in word-medial position, VCV (Rothenberg [1968]: 7.4 for American English, Hindi, and Telugu; Van Hattum and Worth [1967]; Lubker [1973] for American English) and across a word boundary, V+CV (Lubker [1973] for English; Pape et al. [2006] for German) – though no data is available for Spanish. Parallel to this pattern, there is phonological evidence that voiced but not voiceless stops may become prenasalized in medial position (see examples (1), (2), and (5) in Section 3). On the other hand, because voiced stops tend to be spirantized intervocalically in Spanish, nasal leak to support voicing may be rare in this language. Second, listeners were told that the sequence of sounds was a snippet of a longer utterance. Thus, they may have parsed the sequence in different ways, with or without word or phrase boundaries. The issue of a bias towards perceiving a nasal in this context is taken up in Section 5 below.

All the stimuli were normalized for intensity,[9] leaving unmodified the relative amplitude differences between consecutive segments. Care was taken that stimuli did not contain disrupting clicks or chirps. No other manipulation of the tokens was made in order to present the perceptual stimuli as close as possible to natural conditions. Because there was no other modification to the data, and in particular no modification of F_0, the stimuli retained their prosodic characteristics and sounded as if they were incomplete snippets from a longer utterance, rather than

9 For each token, the energy maximum was calculated from an energy contour with a window width of 0.01 seconds including the entire /V(C)CV/ sequence. The energy value for each stimulus was multiplied by the ratio between the mean energy value averaged across all tokens and the energy value for that stimulus. This procedure raised the intensity level of those stimuli that had a lower intensity than average, and lowered the stimuli with a higher-than-average intensity.

being a complete one-word utterance. Subjects were told that the recordings had been excised from longer sentences, and were warned that this might sound somewhat odd.

The perception test consisted of 108 randomized stimuli, i.e., 36 tokens of /b, d/ (3 items × 3 repetitions × 2 segments × 2 levels) and 16 tokens of /p, t/ (2 items × 2 repetitions × 2 segments × 2 levels) for the high and mid delayed nasal closure conditions, 16 tokens of /b, d, p, t/ for the no delayed nasal closure condition (2 items × 2 repetitions × 4 segments), 24 tokens of /b, d/ for the high and mid nasal burst conditions (2 items × 3 repetitions × 2 segments × 2 levels), and 16 tokens of /b, d, p, t/ for the nasal consonant condition (1 item × 4 repetitions × 4 segments).

Let us recap the questions and hypotheses that this experiment set out to test. The general question is whether listeners can detect the nasal leak associated with voicing initiation and reinterpret it as a nasal segment. The specific hypothesis is that, if listeners can in fact hear the nasal leak, they are more likely to report a nasal segment in tokens showing maximum delayed velic closure/largest nasal burst than in tokens with less velum leak (i.e., more nasal responses are expected in high > medium > no nasal leak conditions). A second hypothesis is that, for the same magnitude of nasal leak, listeners are more likely to report a nasal before voiced than voiceless stops because nasal leak is not acoustically detectable without a glottal source.

4.1.2 Test procedure and analyses

A two-choice forced identification task was administered to 34 Spanish undergraduate students, who were asked to identify the stimuli as containing a nasal (VNCV) or not (V(C)CV), and were told that there were approximately even numbers of tokens in each category. Six training trials, three clear examples of nasal stimuli (e.g., /VmbV/, VmpV/), and three clear examples of non-nasal stimuli (i.e., tokens with no velum leak), were presented with feedback for familiarization. The stimuli were presented with E-Prime via headphones. The list was presented in one block and randomized for each listener.

The stimulus was presented over the headphones and the two response options were presented visually on the screen, (1) VNCV in the left-hand area of the screen and (2) VC(C)V on the right for half of the subjects, and the reverse order for the other half. No effect of the right-left location of the responses on the screen was found ($\chi^2[1] = 1.53$, $p = 0.216$), and the results for the two orders were pooled. Subjects identified which of the two sequences they heard by pressing the corresponding key, (1) or (2), of a response pad. Subjects had 3 seconds to respond, and

then the next stimulus was played. Non-responses were removed from the analysis. Reaction time values beyond three standard deviations from the mean were removed from the data set.

The task proved rather difficult as indicated by a rather high error rate. The responses from 6 subjects were discarded because they showed an error rate above 30% in the control items (Nb, Nd, Nt, Nd and b-no, d-no, p-no, t-no), leaving data from 28 subjects for analysis.

Percentage of N identification and reaction time was obtained for each stimulus. In order to handle binary responses (presence vs. absence of N), individual mean responses were calculated as the mean percentage of N responses for each condition for each subject. The mean N responses were submitted to repeated measures ANOVAs. Three-way repeated measures ANOVAs were performed with individual mean N responses as the dependent variable, and Condition (no nasal leak, medium nasal leak, high nasal leak, nasal consonant), Place (labial, coronal), and Voice (voiced /b/, /d/ vs. voiceless /p/, /t/) as fixed factors. Listeners were treated as a random factor. When interaction effects were significant, one-way ANOVAs were performed to examine the effect of each level separately. Note that the delayed nasal closure and the nasal burst patterns for voiced stops only differed in the medium and high conditions; the control no nasal leak and nasal consonant conditions were the same in both patterns.

Because reaction times are positively skewed (i.e., have a lower-bound), a log transformation was performed on the data to obtain a normal distribution. Mixed-model repeated measures ANOVAs were used to analyze the RT data.

4.2 Results

The results of the experiment conducted to evaluate the effect of the acoustic consequences of velum leak on the perception of a nasal consonant are presented in the following sections. The results and statistical analyses focus first on whether listeners could hear a nasal with incremental velum leakage. We hypothesized that listeners would tend to report more nasals with increased velum leak/nasal burst (i.e., high > medium > no nasal leak), and before voiced than voiceless stops. These hypotheses were tested by calculating the number of reported nasals (VNCV responses) out of the total number of possible responses for each experimental condition (high, medium, no velum leak) and the control condition (N consonant) in the voiced and voiceless series. Differences between the delayed nasal closure pattern and the nasal burst pattern were also analyzed. Second, the analysis focuses on the reaction time, and therefore the 'goodness' and/or certainty of the listener for each perceptual stimulus.

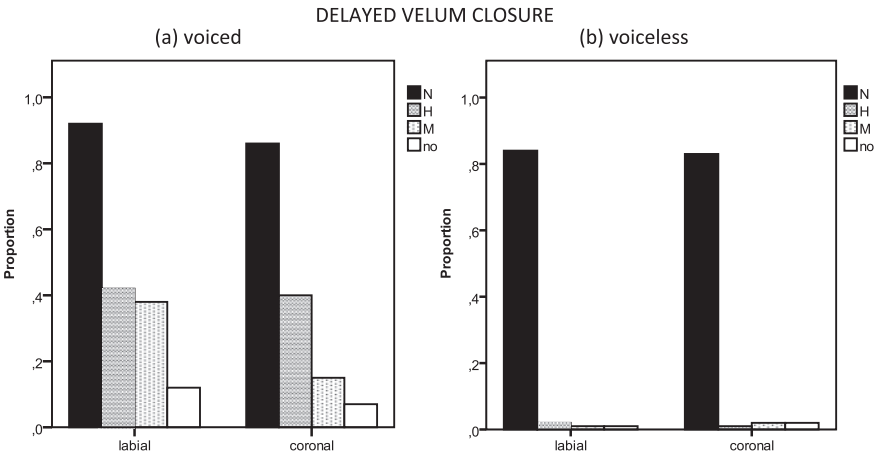

Fig. 2: Percentage of identification of a nasal as a function of decreasing delayed nasal closure (N = nasal consonant, H = high velum leak, M = medium velum leak, no = no velum leak) for labial and dental voiced (a) and voiceless (b) stops.

4.2.1 Perception of a nasal consonant

The results of the identification test are shown in Figures 2 and 3, which show percentage of identification of a nasal as a function of decreasing *delayed nasal closure* for labial and coronal voiced (Figure 2a) and voiceless stops (Figure 2b), and decreasing *nasal burst* for labial and dental voiced stops (Figure 3), averaged across listeners. The bars represent the percentage of VNCV responses in the different conditions (lexical N, high, mid, no nasal leak). (The two extreme categories – lexical N and no nasal leak – are the same for the delayed nasal closure and the nasal burst conditions before a voiced stop; only the high and mid nasal leak categories differ). The complementary values were non-VNCV responses (i.e., VC(C)V). It can be seen that listeners reliably identified the endpoints of the nasal leak continuum as VNCV and non-VNCV, respectively. Thus, listeners identified a nasal in /NC/ tokens (92% for /Nb/, 86% for /Nd/, 84% for /Np/, and 83% for /Nt/) with a trend to hear more nasals in a voiced than in a voiceless context. For the no nasal leak condition – the other endpoint of the continuum – between 7% and 12% of the listeners reported a nasal before a voiced stop vs. only 1–2% before a voiceless stop. Thus voiced stops seem to favor the percept of a nasal consonant both with and without the occurrence of nasal leak. Figures 2 and 3 show a roughly gradual decrease in nasal identification from VNCV to non-VNCV responses for voiced but not for voiceless stops. The figures

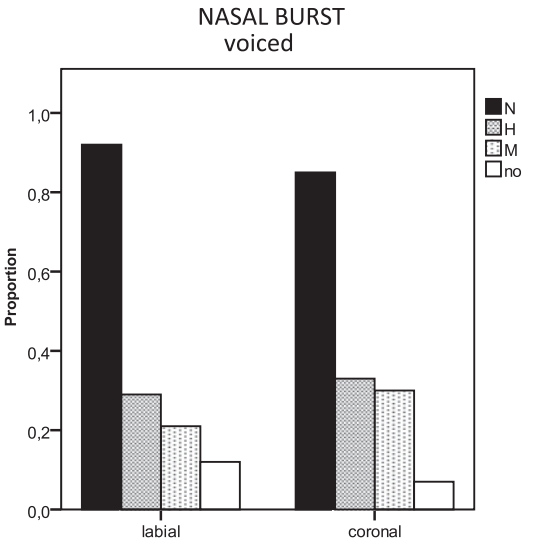

Fig. 3: Percentage of identification of a nasal as a function of decreasing nasal burst (N = nasal consonant, H = high nasal burst, M = medium nasal burst, no = no nasal burst) for labial and dental voiced stops.

show that the predicted effects of nasal leak for a voiced stop were obtained: listeners were more likely to report a nasal consonant with increasing nasal leakage, both in the delayed nasal closure and nasal burst patterns.

These observations are borne out by the statistical analyses. A three-way repeated-measures ANOVA was conducted on the dependent measure of individual mean N-responses (i.e., VNCV responses) for the 'delayed nasal closure' pattern (Figure 2). Only significant differences are reported. Condition (no nasal leak, medium nasal leak, high nasal leak, nasal consonant), Place (labial, coronal), and Voice (voiced /b/, /d/ vs. voiceless /p/, /t/) were within-subject factors. The main effects of Condition [$F(3,25) = 263.48$, $p < .001$, $\eta_p^2 = .969$], Voice [$F(1,27) = 42.28$, $p < .001$, $\eta_p^2 = .610$], and Place [$F(1,27) = 7.40$, $p < .05$, $\eta_p^2 = .215$) were significant. These results indicate that, with incremental changes in delayed nasal closure, a higher number of VNCV responses were obtained in voiced than voiceless stimuli, and in labial than coronal segments.[10] The interaction between

10 The significant effect of Place, with labials showing more nasal identification than coronals, can be interpreted in terms of the former being more conducive to voicing, due to the larger area of compliant tissue (Rothenberg 1968), and hence showing less prenasalization or supporting maneuvers to support voicing than coronals (Solé 2011). Consequently, listeners

Condition and Voice was also significant [$F(3,25) = 20.28$, $p < .001$, $\eta_p^2 = .709$]. The interaction can be observed in Figure 2 in the higher identification of nasals with increasing nasal leak in voiced (a) than voiceless (b) stops. One-way repeated measures ANOVAS examining each voicing class separately (labials and alveolars pooled) show that the effect of Condition was significant in voiced stops [$F(3,53) = 234.78$, $p < .001$, $\eta_p^2 = .930$], and in voiceless stops [$F(3,53) = 202.99$, $p < .001$, $\eta_p^2 = .920$],[11] with a larger effect in the former. Follow-up pair-wise comparisons within the voiced and voiceless stops were conducted to evaluate the differences in perception of nasals (i.e., VNCV responses) among changes in nasal leak (no leak, medium, high, nasal C), with a Bonferroni adjustment of the alpha level to .008 (= .05/6). The results are given in Figure 4a, which shows that, with the exception of high vs. mid nasal leak for labials, nasal identification was significantly higher at each level of delayed nasal closure for voiced tokens (upper right triangle). For voiceless tokens, the only significant difference is between the nasal consonant condition and the rest (lower left triangle). These results indicate a gradual increase in VNCV responses with incremental changes in velum leak for voiced but not voiceless stops. The results suggest that delayed velum closure in voiced stops may in fact result in perceptual identification of a nasal segment.

Comparison of the results for two patterns of nasal leak – delayed nasal closure (Figure 2) and nasal burst (Figure 3) – for voiced stops (voiceless do not show a nasal burst) shows no significant differences. A three-way repeated measures ANOVA with Pattern (delayed nasal closure vs. nasal burst), Condition (N, high, medium, no leak), and Place (labial vs. coronal) shows that the main effect of Condition was significant [$F(3, 25) = 194.10$, $p < .001$, $\eta_p^2 = .959$], but the Pattern [$F(1, 27) = 3.2$, $p > .05$, $\eta_p^2 = .106$] and Place [$F(1, 27) = 2.69$, $p > .05$, $\eta_p^2 = .091$] main effects were not significant. The Place × Pattern interaction [$F(1, 27) = 12.38$, $p < .01$, $\eta_p^2 = .314$] reached significance, indicating that more nasal responses are reported in the delayed nasal closure than in the nasal burst pattern for labials (mean = 0.4614 vs. 0.3918) but not for alveolars (mean = 0.3745 vs. 0.3900), which do not differ in the two patterns. Finally, pair-wise comparisons (Bonferroni) were

may be less likely to attribute prenasalization in labials to voicing maintenance. This interpretation, however, is weakened because no place effect was found in the nasal burst continuum.

11 The d.f. in the repeated-measures one-way ANOVA with the pooled Place has double d.f. ($df = 53$) than the three-way ANOVA ($df = 25$) because, in the one-way ANOVA, the repeated measures for /b/ in the different conditions are now added to the repeated measures for /d/, while the comparison across Place (/b/ vs /d/) was a factor in the three-way ANOVA.

(a) Delayed nasal closure

Labial

voiced

	N	High	Mid	No
N		*	*	*
High	*			*
Mid	*			*
No	*			

voiceless

Coronal

voiced

	N	High	Mid	No
N		*	*	*
High	*		*	*
Mid	*			*
No	*			

voiceless

(b) Nasal burst

Labial

voiced

	N	High	Mid	No
N		*	*	*
High			*	*
Mid				*
No				

Coronal

voiced

	N	High	Mid	No
N		*	*	*
High				*
Mid				*
No				

Fig. 4: Significant differences for the post-hoc pair-wise comparisons (using the Holm's sequential Bonferroni method) for changes in (a) delayed nasal closure and (b) nasal burst. Asterisks indicate significant differences (Bonferroni, $p < .05$). Pair-wise comparisons for voiced tokens are shown in the upper right triangle, and for voiceless tokens (only for delayed nasal closure) in the lower left triangle.

computed to assess differences between incremental conditions in nasal burst for labials and coronals. Figure 4b shows that, except for High vs. Mid burst for coronals, all differences in incremental nasal burst reach significance. No other interaction effects were significant.

In summary, listeners were more likely to report a nasal consonant (i.e., VNCV responses) with increased velum leak, as predicted by hypothesis (1), and in voiced than in voiceless stops, as predicted by hypothesis (2). Differences were found between the two patterns of nasal leak only in labial – with a higher identification of VNCV responses in the delayed nasal closure than the nasal burst continuum – but not coronal voiced stops.

4.2.2 Reaction time analysis

In a second analysis, response time (RT) data for nasal responses (i.e., VNCV) was examined. This restricted the analysis to the two voiced series as the voiceless series had negligible 'VNCV' responses for the test tokens. Reaction time was expected to provide an indication of the listeners' certainty in responding to each stimulus. Based on results showing that low intelligibility slows down response times (e.g., Wilson and Spaulding 2010), we expected to find faster nasal responses not only for control NC tokens, but also for high velum leak than for medium or no velum leak tokens. That is, if listeners are partly relying on nasal leak for VNCV identification, then high velum leak tokens would be considered 'better' exemplars and identified faster. Because reaction times are positively skewed, a log transformation was performed on the data to obtain a normal distribution and carry out the statistical analyses.

The reaction times for voiced stops in the delayed nasal closure and nasal burst series are shown in Figure 5. As expected, the response times for control NC are faster than for the other conditions. The overall pattern of results for both series is that response times get increasingly slower as velum leak diminishes. The crucial observation is that high tokens show faster latencies than medium and no leak tokens. Statistical analyses confirmed this interpretation. Two two-way repeated measures ANOVAs with Condition (N, high, medium, no leak) and Place (labial, coronal) for the delayed nasal closure (1) and nasal burst (2) patterns

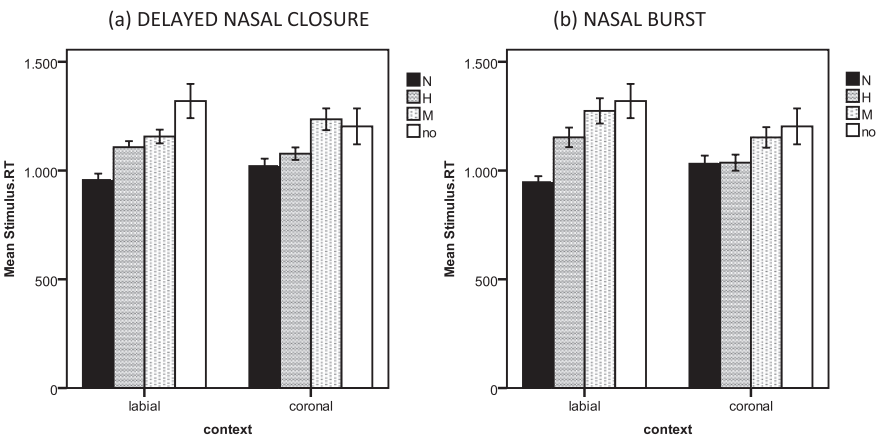

Fig. 5: Mean response time (ms) and standard error for VNCV responses for labial and dental voiced stops as a function of decreasing delayed nasal closure (a) and decreasing nasal burst (b). Error bars: +/–1 SE.

delayed nasal closure

	N	High	Mid	No
N		*	*	*
High			*	*
Mid	*	*		
No	*	*		

nasal burst

Fig. 6: Significant differences for the post-hoc pair-wise comparisons (Bonferroni) for changes in delayed nasal closure in the upper right triangle and nasal burst in the lower left triangle. Asterisks indicate significant differences (Bonferroni, $p < .05$).

separately show that the main effect of Condition was significant in both patterns [$F_1(3, 141) = 32.19$, $p < .001$, $\eta_p^2 = .407$); $F_2(3, 90) = 11.26$, $p < .001$, $\eta_p^2 = .385$], but Place [$F_1(1, 143) = 1.14$, $p > .05$, $\eta_p^2 = .008$; $F_2(1, 92) = 2.85$, $p > .05$, $\eta_p^2 = .048$] and interaction effects [$F_1(3, 141) = 1.97$, $p > .05$, $\eta_p^2 = .040$; $F_2(3, 90) = 2.05$, $p > .05$, $\eta_p^2 = .030$] were not significant. Post-hoc pair-wise comparisons (Bonferroni) for the main effect of Condition (see Figure 6) revealed that for the delayed nasal closure pattern, all contrasts were significant ($p < .05$), except for medium vs. no leak tokens. Thus, each increment in delayed nasal closure tended to yield significantly faster latencies in VNCV identification. Post-hocs for the nasal burst pattern showed that NC and high tokens did not differ in latencies, but both yielded significantly faster responses than medium and no tokens (which did not differ between them). The correlation between degree of nasal leak and processing time suggests that identification of a nasal consonant is affected by the presence of nasal leak.

5 Discussion and conclusions

The identification results show that listeners were more likely to report a nasal consonant with incremental changes in velum leak during voiced but not voiceless stops. The results for voiced stops show that, crucially, high and medium nasal leak stimuli (in both delayed nasal closure and nasal burst patterns) differed in all cases from the two endpoints of the continuum – an intended nasal consonant and absence of a nasal leak (see Figure 4). Thus, these intermediate tokens produced significantly fewer nasal responses than a nasal con-

sonant and more nasal responses than the no leak condition, suggesting that nasal leak is detected by listeners and may trigger nasal identification. Although great care was taken in the selection of the intermediate stimuli and the differences in nasal identification between the high and medium nasal leak conditions were always in the expected direction (high > medium), these differences only reached significance in half of the cases (see Figure 4). It may be that factors other than length of nasal leak (possibly magnitude) can influence the percept of a nasal and, therefore, a more elaborate selection of the tokens is needed. Alternatively, it may be that only duration of nasal leak is relevant but some of the selected tokens were not distinct enough to trigger a different perceptual response.

Another finding was that both types of nasal leak (delayed nasal closure and nasal burst) had similar perceptual consequences for coronals but not for labials, which showed more nasal responses in the delayed nasal leak than the nasal burst pattern. A similar perceptual effect of both patterns would be consistent with Solé and Sprouse's (2011) interpretation of production data on nasal leak. They suggest that occurrence of one or the other pattern largely depends on whether or not the velum is open at the beginning of the utterance. If the velum is open (e.g., during inhalation between phrases or in rest position – note the burst of nasal flow at the end of the utterance in Figure 1), then velum closure tends to be delayed relative to oral closure. If the velum is closed, then it momentarily opens for voiced but not voiceless stops. It is to be expected then that if the occurrence of one or the other pattern depends exclusively on initial conditions, but both have the same functional target (to facilitate voicing) and the same consequence (comparable frequency of occurrence of prevoicing), they would have a similar perceptual effect. The higher nasal identification in the delayed velum closure than in the nasal burst for labials, therefore, remains unexplained (but see note 10).

The results of the reaction time analysis indicate that stimuli with larger velum leak showed shorter latencies to nasal identification than those with smaller nasal leak. Specifically, high and/or medium leak tokens showed shorter latencies than no leak tokens (and longer latencies than nasal tokens), and here high leak stimuli showed shorter response times than medium leak stimuli in 3 out of 4 comparisons (coronals in the delayed nasal closure pattern and labials and coronals in the nasal burst pattern). As RT reflects 'category goodness', tokens with greater nasal leak were considered 'better' exemplars of a nasal than tokens with less leak.

Overall the results suggest that listeners can detect the nasal murmur and may fail to relate it to the initiation of voicing, interpreting instead a nasal segment. The results thus provide empirical evidence of misparsings (or

misperceptions) underlying phonological patterns of emerging nasals. The signal was misparsed when two events – nasal leak and vocal fold vibration – associated with voiced stops were parsed separately, what Ohala (1989, 2012) calls 'hypo-correction' and Ohala and Busà (1995) a 'dissociation' parsing error.

It is surprising, though, to note the large proportion of nasals identified by listeners (Figures 2 and 3), considering the fact that correct perception is the norm and sound change exceptional. We think these results can be accounted for in the following terms. We have reviewed evidence above that shows that nasal leak is relatively commonly used to facilitate voicing in utterance-initial stops. Therefore, Spanish listeners have experience with this adjustment and normally correctly parse nasal leak with voicing in the stop. The large number of reported nasals may reflect failure to attribute nasal leak to voicing in non-utterance initial environments, where nasal leak is not commonly found. In other words, because the test syllable was excised from its utterance-initial context and presented intervocalically, subjects were not able to use their abilities to compensate for the effects of (utterance-initial) boundary (Lindblom and Studdert-Kennedy 1967; Mann and Repp 1980), and the nasal leak associated with voicing initiation became detectable.

Given this scenario, one could argue that the results of the perception experiment do not necessarily demonstrate mis-parsing of an acoustic nasal, but instead that listeners have *correctly* interpreted acoustic nasalization as evidence for a phonological nasal consonant that can, after all, exist in Spanish in medial context, /VNCV/. (Of course, the perceptual tests had to use a medial context because, unlike word onsets, this is a context where nasals can precede stops in Spanish.) There are two points to be made regarding this claim. First, listeners, in interpreting the nasal murmur, which results from ancillary gestures, as a phonological nasal *failed to recover* the signal intended by the speaker, hence the mis-parsing interpretation. Second, it is plausible to suggest that utterance-initial voiced stops in running speech may occur in similar acoustic conditions as those presented in the experiment. The results of the perceptual test indicate that when boundary or contextual information is removed or missed, nasal leak may be interpreted as a phonological nasal. The results therefore provide a baseline picture of how the acoustic effects of nasal leak intended to achieve vocal fold vibration in stops may be reanalyzed when boundary (or context) information is missed by listeners.

Another issue is the perceptibility of nasal leak adjacent to voiced vs. voiceless stops. The results in Figure 2 indicate that nasal leak is more detectable in voiced than in voiceless stops. One could argue that this may be an effect of the slightly lower nasal lag values – reflecting the distribution of values in natural speech – for voiceless than for voiced stops (in the delayed nasal closure pattern).

However, if this were the case, we would expect a smaller but still gradual decrease in nasal identification in the voiceless compared to the voiced series, rather than the virtual absence of /VNCV/ responses seen in Figure 2. Such lack of nasal identification with increasing nasal leak during voiceless stops suggests that, as hypothesized, delayed velum closure is not auditorily detectable without a glottal sound source that resonates in the nasal cavity. A second line of argument supports this interpretation. If nasal leak were equally audible during voiced and voiceless stops, then *fewer* nasals would be reported in the voiced stop condition, as experienced listeners would be likely to attribute such nasal leak to voicing initiation. An audible nasal leak before a voiceless stop, on the other hand, could not be attributed to such adjustment for voicing and would tend to be identified as a nasal segment. Our results point in the opposite direction: more nasals are reported in the voiced stop condition.

Turning to sound change, the results provide evidence that listeners detect the nasal leak and attribute it to the presence of a nasal segment rather than to the initiation of voicing. Phonologists have long suggested that listeners' misperceptions are a source of sound change. Listeners' misperceptions are due to ambiguities in the acoustic signal with respect to articulation. These ambiguities, in turn, may have different sources: (1) a given acoustic pattern may correspond to more than one vocal tract configuration, e.g., American English [r] and [w] are spectrally similar but articulatorily different (McGowan et al. 2004); [θ] and [f] also share spectral characteristics despite their different articulation (Nissen and Fox 2005); (2) contextual or rate-induced variation in the realization of a given segment (or sequence of segments) may correspond very closely to the realization of some other segment, e.g., /du:/ has a falling F2 transition from the alveolar consonant to the back vowel, very similar to the falling F2 for the intended palatal glide in /dju:/); or (3) implementational gestures directed at achieving a certain target goal (e.g., nasal leak to facilitate voicing) may be similar to a different target gesture (velum lowering for a nasal consonant). If the listener incorrectly identifies the articulatory source of the acoustic pattern (e.g., /r/ interpreted as /w/), fails to correct for coarticulatory effects (e.g., attributes the falling F2 in /du:/ to a separate /j/ segment), or interprets the voice-sustaining *strategy* (nasal leak) as a new target *goal* (nasal segment), then when he turns speaker, he produces the incorrectly reconstructed form rather than the original articulation.

In the first two scenarios, the ambiguities are consistent and predictable from (1) the similar frequency response of different articulatory configurations and (2) coarticulatory dynamics, respectively. Listeners are aware of these effects and normally correctly reconstruct the intended signal. Implementational effects, such as nasal leak or oral cavity enlargement to facilitate transglottal flow, on the other hand, are not fully predictable because they vary across and within

speakers (Westbury 1983), across segments of the stop category, and across languages (Solé 2011). It is possible that implementational features, because they are not fully predictable, may be more difficult to attribute to the source (the implementation of voicing) and more likely to be interpreted as a target feature in itself and reproduced as such.

It would be useful as a conclusion to elaborate on how a gesture that was aimed at facilitating voicing initiation may be interpreted as a new target goal. First, the speaker's goal is to produce vocal fold vibration during voiced stops in order to maintain the pronunciation norm (and thus maintain the obstruent voice contrast). Second, in order to achieve this goal in rather challenging aerodynamic conditions, the speaker may use a variety of motor strategies (e.g., velum leakage, larynx lowering, or retroflexion in the case of dento-alveolars [Rothenberg 1968; Westbury 1983; Sprouse et al. 2008]) in much the same way as F2 lowering for /u/ may be achieved by different degrees of tongue dorsum raising or lip rounding ('motor equivalence'; see, for example, Perkell et al. 1993). Note that these strategies, though planned by the speaker, are a means to initiate/sustain vocal fold vibration, not a goal in themselves. This is in line with the view that the regulation of speech involves monitoring not only the articulatory or acoustic-auditory target variable being controlled but also some other functions related to it (for example, altering intraoral pressure in response to changes in the intensity of frication [Moon and Folkins 1991; see also Huber et al. 2004]). The use of different articulatory adjustments to achieve vocal fold vibration is also in keeping with the view that a target gesture involves a 'constellation of gestures' (Browman and Goldstein 1989, 1995), and the recent proposal by Tilsen and Goldstein (2012) that the selection of articulatory gestures associated with a particular segment takes place on an individual basis, and therefore that gestures affiliated with the same segment may fail to occur. Third, an unintended by-product would be that these voice-sustaining strategies become a goal for the speaker, namely, the implementation of prenasalization, implosives, or retroflex apicals as *new target goals* for the speaker.

How does a strategy intended to sustain voicing (i.e., a function related to a goal) become a goal in itself? In exemplar theory, prenasalized exemplars would be added to the probability distribution, shifting the modal value toward the prenasalized pronunciation. Because speakers (and listeners) are sensitive to the distributional patterns in the input data, when sampling from the entire aggregate of exemplars, they are more likely to use prenasalized variants. In Ohala's (1983, 1993) terms, it would be due to a misinterpretation on the part of the listener: the listener reconstructs a new phonological form as a result of drawing wrong inferences about speakers' intentions from the speech signal.

Acknowledgments: This work was supported by grants FFI2010-19206 from the Ministry of Science and Innovation, Spain, and 2009SGR3 of the Generalitat de Catalunya.

References

Baer, Thomas. 1975. Investigation of phonation using excised larynxes. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA. http://dspace.mit.edu/handle/1721.1/21325#files-area

Beddor, Patrice S. 2012. Perception grammars and sound change. In Maria-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change. Production, perception, and social factors,* 37–56. Amsterdam: John Benjamins.

Bell-Berti, Fredericka. 1975. Control of pharyngeal cavity size for English voiced and voiceless stops. *Journal of the Acoustical Society of America* 57(2). 456–461.

Browman, Catherine P., & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6. 201–251.

Browman, Catherine P., & Louis Goldstein. 1995. Dynamics and Articulatory Phonology. In Robert F. Port & Timothy Van Gelder (eds.), *Mind as Motion: Explorations in the dynamics of cognition,* 175–193. Cambridge, MA: The MIT Press.

Brown, Robert. 2001. Waris (Walsa) language. SIL. PNG Language resources. Waris Organised Phonology Data. http://www.pnglanguages.org/pacific/png/pubs/0000361/Waris.pdf

Cohn, Abigail C. 1990. Phonetic and phonological rules of nasalization. *UCLA Working Papers in Phonetics* 76.

Cole, Jennifer, José I. Hualde, & Khalil Iskarous. 1999. Effects of prosodic and segmental context on /g/-lenition in Spanish. In O. Fujimura, B. D. Joseph, & B. Palek (eds.), *Proceedings of the Fourth International Linguistics and Phonetics Conference*, 575–589.

Harrington, Jonathan. 2012. The coarticulatory basis of diachronic high back vowel fronting. In Maria-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change. Production, perception, and social factors,* 103–122. Amsterdam: John Benjamins.

Harrington, Jonathan, Felicitas Kleber, & Ulrich Reubold. 2008. Compensation for coarticulation, /u/-Fronting, and sound change in Standard Southern British: An acoustic and perceptual study. *Journal of the Acoustical Society of America* 123. 2825–2835.

Hayes, Bruce, & Tanya Stivers. 2000. Postnasal voicing. Ms., Dept. of Linguistics, UCLA.

Herbert, Robert K. 1986. *Language universals, markedness theory, and natural phonetic processes*. New York: Mouton de Gruyter.

Hickey, Raymond. 1997. Sound change and typological shift; Initial mutation in Celtic. In Jacek Fisiak (ed.), *Linguistic typology and reconstruction,* 133–182. Berlin: Mouton de Gruyter.

Hirose, Hajime, & Seiji Niimi. 1987. The relationship between glottal opening and the transglottal pressure differences during consonant production. In Thomas Baer, Clarence Sasaki, & Katherine S. Harris (eds.), *Laryngeal function in phonation and respiration*, 381–390. Boston: College Hill.

Huber, Jessica E., Elaine T. Stathopoulos, & Joan E. Sussman. 2004. The control of aerodynamics, acoustics, and perceptual characteristics during speech production. *Journal of the Acoustical Society of America* 116(4). 2345–2353.

Iverson, Gregory K., & Joseph C. Salmons. 1996. Mixtec prenasalization as hypervoicing. *International Journal of American Linguistics* 62(2). 165–175.
Jones, Mark J. 2001. The historical development of a phonological rarity: postnasalized stops in Lancashire dialects. Paper presented at the 9th Manchester Phonology Meeting. http://www.cantab.net/users/m.j.jones.91/MJJ_archive.html
Kent, R. D., & K. L. Moll. 1969. Vocal-tract characteristics of the stop cognates. *Journal of the Acoustical Society of America* 46(6B). 1549–1555.
Kuehn, David P., & Jerald B. Moon. 1998. Velopharyngeal closure force and levator veli palatini activation levels in varying phonetic contexts. *Journal of Speech, Language, and Hearing Research* 41. 51–62.
Laver, John. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.
Li, Fang Kuei. 1977. *A handbook of comparative Tai*. Honolulu: The University Press of Hawaii.
Lindblom, Björn, & Michael Studdert-Kennedy. 1967. On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America* 42(4). 830–843.
Llach, Sílvia. 1999. La neutralització de la sonoritat en els grups consonàntics finals de les formes verbals sense terminació del mallorquí. In *Actes del I Congrés de Fonètica Experimental, Tarragona, 22, 23 i 24 de febrer de 1999*, 241–248. Barcelona: Universitat Rovira i Virgili/Universitat de Barcelona.
Lubker, James F. 1973. Transgottal airflow during stop consonant production. *Journal of the Acoustical Society of America* 53(1). 212–215.
Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
Maddieson, Ian, & Peter Ladefoged. 1993. Phonetics of partially nasal consonants. In Maria K. Huffman & Rena A. Krakow (eds.), *Nasals, nasalization, and the velum,* 251–301. San Diego: Academic Press.
Mann, Virginia A., & Bruno H. Repp. 1980. Influence of vocalic context on oerception of the [ʃ] – [s] distinction. *Perception and Psychophysics* 28. 213–228.
McGowan, Richard, Susan Nittrouer, & Carol J. Manning. 2004. Development of [ɹ] in young, Midwestern, American children. *Journal of the Acoustical Society of America* 115(2). 871–884.
Moon, Jerald B., & John W. Folkins. 1991. The effects of auditory feedback on the regulation of intraoral pressure during speech. *Journal of the Acoustical Society of America* 90. 2992–2999.
Muselhold, A. 1913. *Akustik und Mechanik der menschlichen Stimmorgane*. Berlin: Springer.
Nihalani, Paroo. 1975. Velopharyngeal opening in the formation of voiced stops in Sindhi. *Phonetica* 32. 89–102.
Nissen, Shawn L., & Robert A. Fox. 2005. Acoustic and spectral characteristics of young children's fricative productions: A developmental perspective. *Journal of the Acoustical Society of America* 118(4). 2570–2578.
Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The Production of Speech,* 189–216. New York: Springer-Verlag.
Ohala, John J. 1989. Sound change is drawn from a pool of synchronic variation. In L. E. Breivik & E. H. Jahr (eds.), *Language change: Contributions to the study of its causes*, 173–198. Berlin: Mouton de Gruyter.
Ohala, John J. 1993. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives,* 237–278. London/New York: Longman.

Ohala, John J. 2012. The listeners as a source of sound change: An update. In Maria-Josep Solé, & Daniel Recasens (eds.), *The initiation of sound change. Perception, production and social factors*, 21–36. Amsterdam: John Benjamins.

Ohala, John J., & Maria Grazia Busà. 1995. Nasal loss before voiceless fricatives: A perceptually-based sound change. *Rivista di Linguistica* 7. 125–144.

Ohala, John J., & Manjari Ohala. 1993. The phonetics of nasal phonology: Theorems and data. In Marie K. Huffman & Rena A. Krakow (eds.), *Nasals, nasalization, and the velum*. [Phonetics and Phonology 5], 225–249. San Diego: Academic Press.

Ohala, John J., & Ronald Sprouse. 2003. Effects on speech of introducing aerodynamic perturbations. In Maria-Josep Solé, Daniel Recasens, & Joaquin Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona*. 2913–2916.

Ohala, Manjari, & John J. Ohala. 1991. Nasal epenthesis in Hindi. *Phonetica* 48. 207–220.

Ortega-Llebaria, Marta. 2004. Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/ in Spanish and English. In Timothy L. Face (ed.), *Laboratory approaches to Spanish phonology*, 237–254. Berlin: Mouton de Gruyter.

Pape Daniel, Christine Mooshammer, Phil Hoole, & Susanne Fuchs. 2006. Devoicing of word-initial stops: A consequence of the following vowel? In Jonathan Harrington & Marija Tabain (eds), *Speech production: Models, phonetic processes, and techniques*, 211–226. Psychology Press: New York.

Pater, Joe. 1999. Austronesian nasal substitution and other NC effects. In René Kager, Harry van der Hulst, & Wim Zonneveld (eds.), *The prosody-morphology interface,* 310–343. Cambridge: Cambridge University Press.

Perkell, Joe S., Melanie L. Matthies, Mario A. Svirsky, & Michael I. Jordan. 1993. Trading relations between tongue-body raising and lip rounding production of the vowel /u/: A pilot motor equivalence study. *Journal of the Acoustical Society of America* 93(5). 2948–2961.

Piggot, Glenn. 1992. Variability in feature dependency. *Natural Language and Linguistic Theory* 10. 33–77.

Quigley, Edward C. 2003. Awara phonology. M.A. thesis, University of North Dakota. http://arts-sciences.und.edu/summer-institute-of-linguistics/theses/_files/docs/2003-quigley-edward-c.pdf

Recasens, Daniel, & Aina Espinosa. 2010. A perceptual analysis of the articulatory and acoustic factors triggering dark /l/ vocalization. In Daniel Recasens, Fernando Sánchez Miret, & Kenneth Wireback (eds.), *Experimental phonetics and sound change*, 71–82. München: Lincom Europa.

Recasens, Daniel, Aina Espinosa, & Antoni Solanas. 2004. Underlying voicing in Majorcan Catalan word-final stop-liquid clusters. *Phonetica* 61(2–3). 95–118.

Rothenberg, M. 1968. *The breath-stream dynamics of simple-released-plosive production*. Basel: Karger.

Slifka, Janet. 2000. Respiratory constraints on speech production at prosodic boundaries. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Smith, Geoff P. 2002. *Growing up with Tok Pisin: Contact, creolization and change in Papua New Guinea's national language*. London: Battlebridge Publications.

Solé, Maria-Josep. 2009. Acoustic and aerodynamic factors in the interaction of features. The case of nasality and voicing. In Marina Vigario, Sonia Frota, & Maria Joao Freitas (eds.),

Phonetics and Phonology. Interactions and Interrelations, 205–234. Amsterdam: John Benjamins.

Solé, Maria-Josep. 2011. Articulatory adjustments in initial voiced stops in Spanish, French and English. In Wai-Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, 1878–1881. Hong Kong: China.

Solé, Maria-Josep, & Ronald Sprouse. 2011. Voice-initiating gestures in Spanish: Prenasalization. In Wai-Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, 72–75. Hong Kong: China.

Solé, Maria-Josep, Ronald Sprouse, & John J. Ohala. 2008. Voicing control and nasalization. *Laboratory Phonology 11* (reviewed abstracts), Victoria, New Zealand. 127–128.

Soler, Antonia, & Joaquín Romero. 1999. The role of duration in stop lenition in Spanish. *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco. 483–486.

Sprouse, Ronald. 2008. A discrete-time aerodynamic model of the vocal tract. *Proceedings of the 8th International Speech Production Seminar*, Strasbourg, France. 337–340.

Sprouse, Ronald, Maria-Josep Solé, & John J. Ohala. 2008. Oral cavity enlargement in retroflex stop. *Proceedings of the 8th International Speech Production Seminar*, Strasbourg, France. 425–428.

Tilsen, Sam, & Louis Goldstein. 2012. Articulatory gestures are individually selected in production. *Journal of Phonetics* 40(6). 764–779.

Torreblanca, Máximo. 1979. Un rasgo fonológico de la lengua española. *Hispanic Review* 47(4). 455–468.

Van Hattum, R. J., & J. W. Worth. 1967. Air flow rates in normal speakers. *Cleft Palate Journal* 4. 137–147.

Ward, Ida C. 1939. A short phonetic study of Wolof (Jolof) as spoken in the Gambia and in Senegal. *Africa* 12. 320–334.

Westbury, John R. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* 73. 1322–1336.

Westbury, John R., & Patricia A. Keating. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145–166.

Wiebe, Brent. 1997. Bola (Bola-Bakovi) Language. SIL. PNG Language resources. Bola Organised Phonology Data. http://www.pnglanguages.org/pacific/png/pubs/0000069/Bola.pdf

Wilson, Erin O., & Tammie J. Spaulding. 2010. Effects of noise and speech intelligibility on listener comprehension and processing time of Korean-accented English. *Journal of Speech and Hearing Research* 53(6). 1543–1554.

Yamane-Tanaka, Noriko. 2005. The implicational distribution of prenasalized stops in Japanese. In Jeroen van de Weijer, Kensuke Nanjo, & Tetsuo Nishihara (eds.), *Voicing in Japanese,* 123–156. New York: de Gruyter.

Yanagihara, Naoki, & Charlene Hyde. 1966. An aerodynamic study of the articulatory mechanism in the production of bilabial stop consonants. *Studia Phonologica* 4. 70–80.