DE GRUYTER MOUTON Laboratory Phonology 2014; 5(4): 455–488

Mary Stevens* and Ulrich Reubold

Pre-aspiration, quantity, and sound change

Abstract: Geminate voiceless stops /pː tː kː/ have been recently found to show optional pre-aspiration under certain circumstances in spontaneous and read Italian speech. This paper investigates the impact of pre-aspiration on the perception and production of contrastive quantity, e.g., *fato* 'fate' vs. *fatto* 'done'. It tests the hypothesis that synchronic variability involving pre-aspiration, together with concomitant stop closure shortening, may be setting in motion a sound change in Italian ultimately leading to de-gemination, i.e., /pː tː kː/ > [ʰp ʰt ʰk] > /p t k/. The proposed sound change would be perceptually driven (Ohala 1981, 1993) and comes about via listener association of pre-aspiration with the preceding vowel rather than the oral closure. The hypothesis is only partially supported by the experimental results. Perception data show that Italians perceive pre-aspirated stops as shorter than plain stops of analogous overall duration. However, production data show that pre-aspiration in Italian does not in fact involve concomitant oral closure shortening.

Keywords: geminate, pre-aspiration, sound change, Italian, acoustic

DOI 10.1515/lp-2014-0015

1 Introduction

1.1 Consonant length in Romance and in Italian

Standard Italian and most centro-southern varieties are exceptional in having maintained geminates from late Latin, e.g., UACCA(M) > Italian *vacca* /vakːa/ 'cow'. Contrastive length has been lost almost everywhere else in Romance languages, e.g., Spanish *vaca* /baka/, French *vache* /vaʃ/, and Romanian *vacă* /vakə/, often accompanied by lenition of intervocalic /p t k/ (cf. Loporcaro 2011: 150–153). In Italian, the primary phonetic cue to geminates is increased consonant constriction duration (cf. Esposito and Di Benedetto 1999; Pickett et al. 1999;

***Corresponding author: Mary Stevens:** Institute for Phonetics and Speech Processing, Munich.
E-mail: mes@phonetik.uni-muenchen.de
Ulrich Reubold: reubold@phonetik.uni-muenchen.de

Payne 2005; Hualde and Nadeu 2011). The preceding vowel plays a secondary role, whereby stressed vowels are predictably short preceding geminates *fatto* [ˈfatːo] 'done' and long preceding singletons *fato* [ˈfaːto] 'fate'. The secondary role of vowel duration in both production and perception has been confirmed experimentally (Esposito and Di Benedetto 1999; Pickett et al. 1999). There is some evidence of overlap between singletons and geminates in Italian when they are measured in terms of consonant duration alone (Payne 2005: 172; Hualde and Nadeu 2011: 223), but, importantly, Pickett et al. (1999) show that a third cue to consonant length, namely the C/V ratio, remains robust even across different speaking rates. Specifically, Pickett et al. found that the threshold value of 1.0 reliably distinguished post-tonic singletons and geminates in both production and perception: geminate consonants must be at least as long as the preceding vowel (i.e., C/V > 1 = /Cː/) whereas singleton consonants are shorter (i.e., C/V < 1 = /C/). For their study Pickett et al. measured voiceless stops /p t pː tː/ in terms of visible closure duration on acoustic displays, not including VOT or any post-aspiration. This is understandable given that descriptive sources all agree that voiceless stops are unaspirated in standard Italian (e.g., Schmid 1999; Bertinetto and Loporcaro 2005; Krämer 2009). Substantial post-aspiration (as distinct from pre-aspiration, the focus of the present study) is occasionally visible and audible in spoken standard Italian (cf. Stevens and Hajek 2010). Voiceless stops are also reported to be post-aspirated in a very small number of regional varieties (e.g., Rohlfs 1966; Mele 2009). In the dialect spoken in San Giovanni in Fiore, Calabria, post-aspiration is reported to be confined to geminates (Mele 2009: 153) but is not considered to be a phonetic cue to consonant length by this author.

1.2 Pre-aspiration and sound change

Recent acoustic phonetic investigations by the first author show evidence of a novel pronunciation for Italian whereby some native speakers optionally produce voiceless geminate stops with aspiration before consonant closure, i.e., pre-aspiration. Other acoustic phonetic studies on Italian involving voiceless geminate stops have not reported pre-aspiration (cf. Esposito and Di Benedetto 1999; Pickett et al. 1999; Payne 2005; Hualde & Nadeu 2011), but in a cross-linguistic study of glottal flow during stop production involving inverse filtering Ní Chasaide and Gobl (1993) reported breathy voice in the transition into voiceless geminate stops for Italian. These authors describe relatively early glottal abduction resulting in breathy voice preceding voiceless stops as a language-specific pattern for their five (Florentine) Italian speakers. This is essentially the inter-gestural timing for pre-aspirated stops, for which glottal abduction must precede oral closure (cf., e.g., Ladefoged and Maddieson 1996: 71–72). Phonologically contrastive pre-aspirated stops are described as rare in the world's languages (cf. Silverman 2003) and are associated mainly with Icelandic and other languages spoken in far northwestern Europe (cf., e.g., Pind [1996] on Icelandic; Nance and Stuart-Smith [2013] on Scottish Gaelic). Optional pre-aspiration in the production of voiceless stops is a widespread tendency in dialects spoken throughout Scandinavia (Helgason 2002). In terms of the Romance languages, pre-aspirated stops are found in several varieties of Spanish as a result of debuccalization of /s/ in coda position, e.g., *estado* [e^{h}taðo] 'state' (e.g., Romero 1995; Torreira 2012; Ruch and Harrington 2014). Figure 1 shows two separate productions of the word *gatto* 'cat' drawn from the CLIPS corpus of spoken Italian (www.clips.unina.it), one involving a plain unaspirated stop and the other with visible and audible pre-aspiration [ga^{h}tːo].

Preliminary acoustic evidence for Italian suggests pre-aspirated realisations of /pː tː kː/ are not associated with any particular regional pronunciation (Stevens 2010, 2011), in contrast to, e.g., Swedish (Wretling et al. 2003) or Norwegian (van Dommelen 1999, 2011). This pre-aspiration also appears not to be confined to any particular speaking style, as it has been documented in read speech recorded in experimental settings (Stevens 2010, 2011 on the CLIPS corpus) as well as in spontaneous speech (Stevens 2012 on Sienese Italian). Some Italian speakers tend to pre-aspirate more than others (cf. Stevens 2012: 224), and structural factors such as consonant place and lexical stress also appear to play a role in the distribution and duration of pre-aspiration (Stevens 2011). In both of these respects, pre-aspiration in Italian broadly resembles phonetic descriptions of pre-aspiration in other languages in which it is an optional feature (e.g., Foulkes et al. 1999 on Newcastle English; Helgason 2004; Helgason and Ringen 2008 on Central Standard Swedish). Pre-aspiration in Italian is unexpected in the sense that Italian has neither contrastive aspiration nor /h/; this conflicts with Silverman's (1997: 7) suggestion that pre-aspiration should only develop in languages that already have contrastive post-aspiration.

Novel synchronic variants such as pre-aspiration in Italian can become permanently contrastive in a language; this idea is well established in the sound change literature (e.g., Ohala 1981, 1993; Beckman et al. 1992; Hura et al. 1992; Hansson 2008; Harrington 2012; and many others). Moreover, the specific sound change from geminates to pre-aspirated stops has been documented to have taken place in other languages, for example, voiceless geminates inherited from Old Norse are now realized with contrastive pre-aspiration in Icelandic, e.g., *kappi* [k^{h}ahpɪ] 'hero' (example from Thráinsson 1978: 6, cf. Kümmel 2007: 172). The motivations and phonetic mechanisms of this sound change, however, are

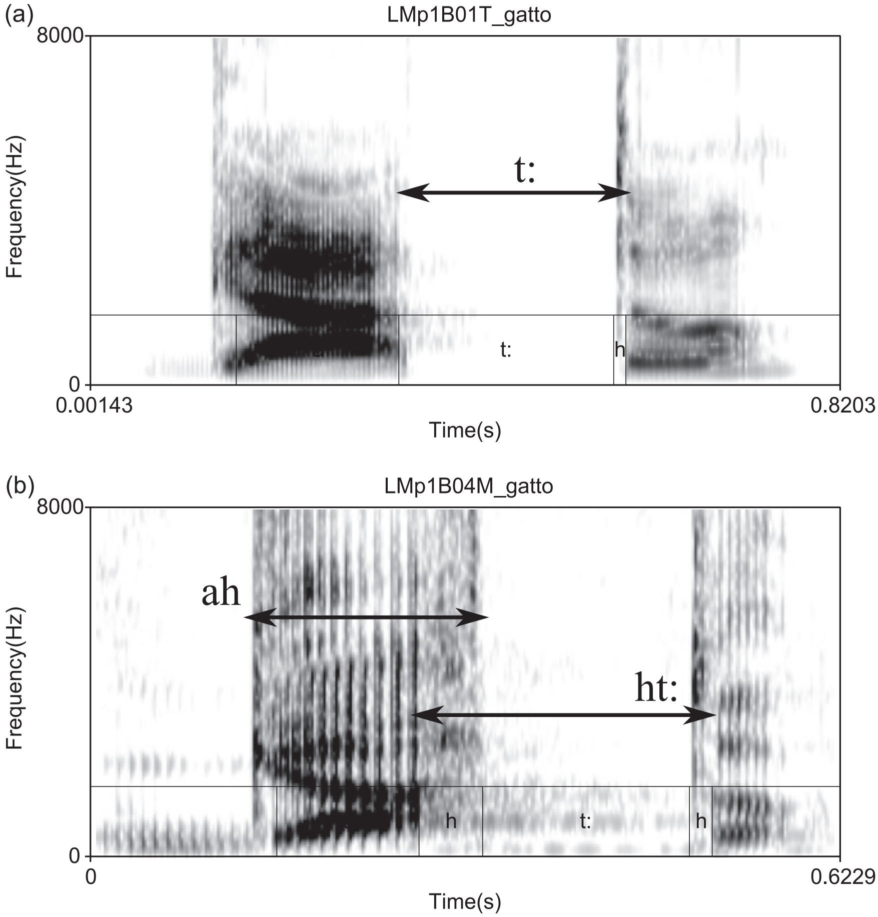

Fig. 1: *Gatto* 'cat' produced with a voiceless unaspirated stop [tː] by one speaker from Turin (a) and with pre-aspiration by a speaker from Milan (b). Arrows indicate possible parsings of pre-aspiration with the vowel or with oral closure (see text).

not well understood and have not yet been the subject of an experimental phonetic investigation. Blevins and Garrett (1993) suggest that pre-aspiration originates in slightly mistimed oral and glottal gestures in the production of voiceless geminate stops. In other words, for these authors pre-aspiration is, at its origins, not under the control of the speaker (cf. Solé [2007] on controlled vs. mechanical phonetic events). Most other sources instead assume pre-aspiration has its origins in a deliberate gesture intended to maintain, e.g., contrastive voicing (Ní Chasaide 1985) or length (Ó Baoill 1980). The development of pre-aspiration in

/pː tː kː/ is nearly always linked to closure shortening, although sources disagree as to whether pre-aspiration triggers closure shortening or vice versa (cf. Ní Chasaide 1985 for discussion). Thráinsson (1978: 35) suggests that for pre-aspiration to develop in geminates, a diachronic rule “has to wipe out or delete the set of supralaryngeal features for the first half of the stop and leave only the laryngeal feature specifications”. The assumption here is that pre-aspiration with concomitant closure shortening involves less supra-laryngeal effort for the speaker and preserves the duration of intervocalic voicelessness, which is assumed to be an important cue to /pː tː kː/ for listeners (cf. Stevens [2012: 249] on Italian). The central aim of this paper is to test these two assumptions about the development of pre-aspiration in /pː tː kː/, i.e., whether

- pre-aspiration involves concomitant shortening of oral closure in production (cf. Section 1.3).
- listeners show a trading relation in perception between pre-aspiration and closure duration (cf. Section 1.4).

Italian listeners attend primarily to the duration of oral closure in making perceptual judgements about phonological quantity (cf. Section 1.1). This means that when /pː tː kː/ are produced with pre-aspiration and concomitant oral closure shortening, they might be more likely to be perceived by Italian listeners as short /p t k/. This misperception of pre-aspirated /pː tː kː/ as short could ultimately lead to de-gemination in contemporary Italian – a novel hypothesis which is outlined at Section 1.5. First, however, pre-aspiration in relation to the production and perception of /pː tː kː/ is addressed in more detail in the following two sections, respectively.

1.3 Pre-aspiration in the production of /pː tː kː/

Available acoustic evidence for present-day Italian is consistent with the first hypothesis just listed above, i.e., that pre-aspiration involves concomitant shortening of oral closure in production. Duration measurements for vowel + /pː tː kː/ sequences in isolated words (e.g., /akː/ in *macchina* /ˈmakːina/ ‘car’) are listed in Table 1 according to whether they were produced with (V^hC) or without (VCː) pre-aspiration. These data were drawn from the CLIPS corpus of spoken Italian (Savy and Cutugno 2009). The fourth column shows that mean closure duration is 30 ms shorter when it follows a portion of pre-aspiration.

The patterns in Table 1 are similar to results reported for pre-aspirated stops in Icelandic, where Ladefoged and Maddieson (1996: 70) note that “the duration of the pre-aspiration and stop closure together in hp, ht, etc. is about equal to the

Table 1: Mean duration (in ms) and standard deviations (in parentheses) for the component parts of vowel + /pː tː kː/ sequences with (V^hC) and without (VCː) pre-aspiration. Adapted from Stevens (2011: 25).

	Vowel (V)	Pre-aspiration	Consonant (C)		*n* tokens
			Closure	Release	
V^hC	98.1 (28)	48.4 (19)	154.4 (48)	52.8 (28)	252
VCː	101.5 (38)	0	183.3 (45)	44.3 (38)	683

duration of the stop closure itself in unaspirated stops".[1] Comparisons between pre-aspirated and plain /pː tː kː/ in spontaneous speech (Stevens 2012, six speakers) also report similar patterns in terms of the impact of pre-aspiration on closure duration. Therefore at present there appears to be free variation for Italian speakers between producing geminates with a long oral closure or with pre-aspiration and a somewhat shorter oral closure. However, available acoustic evidence for pre-aspiration in Italian is based on pre-existing corpora for which the potential impact of speaker identity, word length, and place of articulation and identity of the surrounding segments could not always be controlled. Whether pre-aspiration necessarily involves concomitant oral closure shortening in the production of voiceless geminate stops remains to be tested.

1.4 Listener perception of pre-aspirated /pː tː kː/

This paper also tests the second hypothesis listed at Section 1.2 above, i.e., whether Italian listeners will show a trading relation between pre-aspiration and closure duration in perception. It does so within the terms of Ohala's (e.g., 1981, 1993) model, which is based on the idea that listeners can select novel synchronic variants for sound change. Speech perception involves parsing auditory information to its source (e.g., McMurray and Farris-Trimble 2012: 389–390 and references therein). Ohala argues that listeners are normally very good at coping with ambiguities in the speech signal and factoring out distortions due to, e.g., background noise or coarticulation. Nonetheless, he suggests that they occasionally make errors, and that these listener errors are the source of sound change. Seen in terms of Ohala's model, the question here is whether Italian listeners will show

1 In Icelandic, historical /bː dː gː/ are realized as long voiceless unaspirated stops. Pre-aspirated and unaspirated stops are both possible realizations of /pː tː kː/ in our Italian data.

a trading relationship between pre-aspiration and oral closure (ensuring the stability of phonological length for /pː tː kː/), or whether listener misparsing of pre-aspiration might bring about sound change in Italian.

The first scenario in terms of Ohala's model is that listeners will attribute pre-aspiration to a distortion in the production of /pː tː kː/ and will correct for its presence in the auditory signal. As long as listeners associate pre-aspiration with (shorter) oral closure in perception, they should still perceive pre-aspirated pronunciations as long.

(1a) speaker produces: [faʰto]
↓
listener hears: [fatːo]
↓
listener maps to: *fatto* 'done'
↓
listener-as-speaker produces: [fatːo]

In (1a) listeners parse pre-aspiration with the oral closure for /tː/. (The illustrations in (1a), (1b), and (1c) are adapted from Ohala's (1981: 182) original depictions of sound change involving /u/-fronting.) In doing so, they essentially correct for pre-aspiration, hearing [tː], and their own production (cf. 'listener-as-speaker') thus remains unchanged. Under this scenario pre-aspirated productions of /pː tː kː/ are not expected to bring about any change to pronunciation norms or to the system of phonological contrasts (cf. Janda and Joseph 2003: 17–18 on speaker innovation vs. language change). On the other hand, Ohala's model allows for a second scenario in which the pronunciation norm for geminates could change. This would be the case if the listener were to detect pre-aspiration (still parsing it with the oral closure for /tː/). Ohala and Busà (1994: 13–14), for example, suggest that /s/-aspiration comes about via listener reinterpretation of breathy voice in the transition to /s/ as an intended event.

This scenario is depicted for pre-aspirated stops in (1b), in which the listener detects pre-aspiration, parses it with oral closure for /tː/, and produces /tː/ as [ʰt] thereafter.

(1b) speaker produces: [faʰto]
↓
listener hears: [faʰto]
↓
listener maps to: *fatto* 'done'
↓
listener-as-speaker produces: [faʰto]

There is no change to the phonemic contrast between /t/ and /tː/ in (1b); it is the pronunciation norm for /tː/ that changes such that part of the long closure is replaced with pre-aspiration. The scenario in (1b) is perhaps the mechanism by which historical voiceless geminates came to be realised with contrastive pre-aspiration in Icelandic. In both (1a) and (1b) there is a match between production and perception in the sense that both show a trading relationship between pre-aspiration and oral closure for /tː/.[2] This trading relation ensures the stability of phonological length over time, despite the novel pronunciation of /pː tː kː/ with pre-aspiration.

Ohala's model provides for an alternative third scenario as well, whereby the listener "fails to parse or associate a given perturbation or variation in the speech signal with the conditioning environment" (1993: 264). The prediction for pre-aspiration in Italian is that native listeners might fail to associate pre-aspiration (the perturbation or variation) with the oral closure for the voiceless geminate (the conditioning environment). As described below, this parsing strategy would favour mapping pre-aspirated pronunciations of /pː tː kː/ to singleton /p t k/ in perception.

1.5 A new proposal: de-gemination via pre-aspiration

To distinguish singletons from geminates, Italian listeners attend primarily to oral closure duration, alone and relative to the duration of the preceding vowel (cf. Section 1.1). Available acoustic evidence (cf. Table 1) shows that mean oral closure is 30 ms shorter when geminate /pː tː kː/ are produced with pre-aspiration. This means that Italian listeners would be more likely to map pre-aspirated pronunciations to short stops /p t k/ – unless of course they parse pre-aspiration with oral closure as just outlined at (1a) and (1b). Note that Ohala's (1993: 246) model predicts that novel perceptual variants form the model for that listener's own speech – thus the listener-turned-speaker produces *fatto* with short oral closure thereafter. The third scenario for listener perception of pre-aspirated stops is illustrated for *fatto/fato* in (1c).

2 The scenario in (1b) could also be captured within the terms of Lindblom et al.'s (1995) model of sound change, which differs from Ohala's in that it contains the idea that listeners occasionally pay special attention to *how* something is said and can choose to reproduce a novel variant in their own speech (cf., e.g., Harrington [2012] on these two models of sound change).

(1c) speaker produces: [faʰto]
↓
listener hears: [faːto]
↓
listener maps to: *fato* 'fate'
↓
listener-as-speaker produces: [faːto]

Why would Italian listeners fail to parse pre-aspiration with the oral closure? There are two main reasons. First, pre-aspiration is hard to hear (Bladon 1986). Native listeners of pre-aspirating languages can detect the onset of pre-aspiration and parse it appropriately (e.g., Ní Chasaide 1985; Pind 1996; Wretling et al. 2003; Helgason 2004), but Helgason (2002: 34) suggests that pre-aspiration is otherwise a subtle feature that can go unnoticed in auditory analyses. Second, the relationship between pre-aspiration and quantity, particularly in perception, is not straightforward. For example, whether native Swedish listeners parse pre-aspiration with consonant closure or not depends on their dialect background. VʰC stimuli with pre-aspiration occupying 28% of the stimulus are more likely to be perceived by listeners from Vemdalen as containing a long consonant, e.g., /takː/ 'thanks' but by listeners from Arjeplog as containing a long vowel, e.g., /taːk/ 'roof' (Wretling et al. 2003). These authors found that the dialect-specific parsings of pre-aspiration with the consonant (Vemdalen) or the vowel (Arjeplog) were also matched in production. Based on production data for Norwegian, van Dommelen (1999) instead suggests that pre-aspiration should be interpreted as an independent segment in that language, i.e., a feature associated neither with the consonant nor with the preceding vowel. Thus there are three different patterns in terms of the relationship between pre-aspiration and quantity contrasts. The present study focuses specifically on whether Italian listeners parse pre-aspiration with the (shorter) oral closure. As we have just seen in (1a), (1b), and (1c), it is this issue that is likely to directly impact the stability of the phonological length contrast in Italian. If Italian listeners succeed in parsing pre-aspiration with oral closure, they will map pre-aspirated stops to geminates; if they fail to do so, Italian listeners will map pre-aspirated stops to singleton /p t k/. Within the latter scenario, the secondary issue arises as to whether Italian listeners associate pre-aspiration with the vowel (like Arjeplog listeners), or as an independent segment (as in Norwegian). The results of the perception experiment are discussed in relation to this secondary issue in Section 2.2.

The three different patterns in terms of the relationship between pre-aspiration and quantity contrasts across languages likely come about because of the perceptual ambiguity of pre-aspiration itself. Archetypal pre-aspiration [h] shares spectral properties with the preceding vowel as well as the consonant: it shows continued spectral activity in the formant regions of the preceding vowel and shares voicelessness with the stop portion. Pre-aspiration is also highly variable and can, for example, “appear either as a period of high-frequency frication before the voiceless stop ‘gap’, or as a breathy continuation of a preceding vowel” (Foulkes et al. 1999: 1626 on optional pre-aspiration in Newcastle English). There is descriptive evidence that pre-aspiration can alternate with vowel lengthening, both synchronically – e.g., Tarascan [ˈtsɨʰkuni̥] ~ [ˈtsɨːkuni̥] ‘to drop from one’s hand’ (Foster 1969 in Silverman 2003: 578) – and diachronically. Silverman (2003) attributes this alternation to the fact that pre-aspiration is hard to hear; the onset of pre-aspiration is presumably more difficult for listeners to detect than the abrupt transition from a vowel into voiceless closure for plain stops (e.g., Hawkins 2010: 66 on the auditory significance of abrupt boundaries). Notably, Italian listeners might have special difficulty with the presence of pre-aspiration in the auditory signal. Like most Romance languages, Italian has neither phonemic /h/ nor contrastive (post-)aspiration. Mielke (2003) reports that listener detection of glottal [h] is dependent on whether the listener’s native phonemic inventory contains /h/ and in which phonotactic positions it can occur. Italian listeners should therefore be relatively poor at detecting pre-aspiration preceding oral closure, and might therefore tend not to parse the two together.[3]

As depicted in (1c), failure to parse pre-aspiration with oral closure could eventually lead to de-gemination. This pathway to de-gemination does not account for de-gemination in other consonant types that would be unaffected by pre-aspiration, such as sonorants or voiced stops. However, shorter oral closure for /pː tː kː/ could, over time, serve to weaken this cue in the production and perception of quantity for other consonant types, such that they might eventually also undergo de-gemination. Some indirect evidence in support of the idea that length may be less stable for voiceless stops is that in contemporary Western Andalusian Spanish, short voiceless stop variants coexist alongside long sonorants (Torreira 2012).[4]

3 In contrast to Italian, Scandinavian languages, e.g., Swedish, Danish, and Icelandic, all have phonemic /h/. This structural difference could perhaps help to explain the development of contrastive pre-aspiration in several Scandinavian varieties.

4 The sound change in progress proposed here for contemporary Italian can potentially be linked to diachronic change in the past. Historical /pː tː kː/ have already been lost in most other Romance varieties (Section 1.1), which, notably, also typically lack /h/. Historical de-gemination has been described as a gradual process in terms of its geographical spread throughout the Romance-speaking area (e.g., Loporcaro 2011: 151) or within a particular variety due to structural factors such as lexical stress (e.g., Rohlfs 1966: 323; Loporcaro 2011: 153). The idea that historical de-gemination may have been phonetically gradual and may have involved an intermediate

In sum, this paper aims to test whether a sound change involving degemination of /pː tː kː/ via pre-aspiration is underway in contemporary Italian. The proposed sound change is perceptually driven and can be seen within the terms of Ohala's model, which links synchronic variation to permanent sound change. The first stage in time involves free variation for speakers between producing /pː tː kː/ with long closure or with shorter closure and pre-aspiration (present-day Italian appears to correspond to this first stage). As long as listeners parse pre-aspiration with oral closure, phonological consonant length should remain stable. In the second stage, listeners fail to parse pre-aspiration with (shorter) oral closure for geminate /pː tː kː/ and thus perceive pre-aspirated pronunciations as short. This model rests on two hypotheses: (1) that pre-aspiration involves concomitant oral closure shortening in production and (2) that listeners do not parse pre-aspiration with the oral closure. These two hypotheses are tested separately in two experiments, described in turn below. The production and perception data are then discussed with special reference to individual variability. Although phonetic studies tend to concentrate on group-level patterns, recent experimental approaches to sound change converge on the idea that differences *between* individuals might be the key to understanding why some synchronic variability becomes sound change, and how change might spread from one individual's grammar to the wider community (e.g., Beddor 2009; Grosvald and Corina 2012; Yu 2013; Stevens and Harrington 2014). With this in mind, individual differences between the production and perception of pre-aspiration are investigated in Section 4.4.

2 Perception experiment

2.1 Method

Listener parsing of pre-aspiration was tested by comparing responses to two 10-step *fatto-fato* continua, one involving plain stops [faːto]…[fatːo], and the other in which part of the oral closure was replaced with pre-aspiration [faʰto]…[faʰtːo]. Importantly, oral closure was *replaced* by pre-aspiration in the perception stimuli in order to represent /pː tː kː/ as they occur in the ambient language, which has

stage of pre-aspiration for voiceless stops differs from the traditional assumption that it was a one-step process in Romance (e.g., Rohlfs 1966; Loporcaro 1996, 2011; Giannelli and Cravens 1997).

been argued to involve concomitant closure shortening (cf., e.g., Table 1). In both continua, the consonant-to-vowel (C/V) ratio decreased from 2.1 (= *fatto* [atː]) to 0.6 (= *fato* [aːt]). Endpoint values were based on Pickett et al.'s (1999: 145) mean values for three speakers and were chosen to be well above and below their suggested threshold of 1.0 for post-tonic geminates. In calculating the C/V ratio for the pre-aspirated [faʰto]...[faʰtːo] stimuli, pre-aspiration was included as part of the overall duration of the consonant (i.e., C = pre-aspiration + closure; V = vowel duration). The hypothesis is that listeners will fail to parse pre-aspiration with the (shorter) oral closure and will therefore give more 'fato' responses for the pre-aspirated stimuli.

2.1.1 Participants

Participants for this experiment were recruited locally in Prato, Tuscany, which lies near Florence. Contemporary standard Italian is spoken with distinct, readily identifiable regional accents (Bertinetto and Loporcaro 2005; Savy and Cutugno 2009), and it is necessary to control for the regional origin of participants as far as possible. For example Northern varieties lack geminates and this feature can carry over into standard pronunciation. Prato Italian sounds very close to the (abstract) notion of standard Italian except that intervocalic singleton /p t k/ are optionally realized as voiceless fricatives due to the *Gorgia toscana*, e.g., *prato* > [praːθo ~ praːho] 'Prato (town name)/field'); cf. Dalcher [2008] on the *Gorgia toscana* in Florentine Italian). Allophonic weakening of intervocalic /p t k/ is typical of almost all non-northern varieties (Giannelli and Cravens 1997: 35), complicating phonetic investigations into stop contrasts (see, e.g., Hualde and Nadeu 2011 on Rome Italian).[5] Due to the *Gorgia toscana*, which is optional but highly salient, Prato listeners have some linguistic experience with glottal [h] and may be better at detecting pre-aspiration in the auditory signal than listeners from other parts of Italy. The *Gorgia toscana* does not directly explain the occurrence of pre-aspirated /pː tː kː/ in Tuscan varieties, however, because speakers from all over the Italian peninsula show this tendency (Stevens 2011).

Results for the perception experiment are based on 15 participants, 4 males and 11 females. All participants were born and raised in Prato, aged between 17

5 The optionality of the *Gorgia toscana* means that short /p t k/ can be realized as [p t k]; cf. Stevens (2012: 73–76) on the potential for phonetic overlap between singletons and geminates in Tuscan varieties.

and 30 years, and studying at the local high school or university.[6] They were all self-reported native speakers of Italian with some experience of a language other than Italian (typically English or French) at school. None reported any difficulties with speech or hearing. Participants were naïve as to the purpose of the experiment and were paid for their participation.

2.1.2 Stimuli and task

To create the stimuli, one *gatto* token was extracted from the CLIPS word list data. In this particular *gatto* repetition (seen in the lower part of Figure 1), the geminate /tː/ was produced with 56 ms of pre-aspiration. The duration of the pre-aspiration portion was adjusted slightly to match the mean of 48 ms reported in an earlier study (Stevens 2011, cf. Table 1 above), and word-initial /g/ was replaced with /f/. This [faʰtːo] stimulus was used as the geminate endpoint for the pre-aspirated continuum. The geminate endpoint for the plain continuum was created by replacing pre-aspiration with 48 ms silence to indicate a long oral constriction. Only closure duration varied across these two stimuli according to the presence [faʰtːo] or absence [fatːo] of pre-aspiration; vowel duration was kept constant. Table 2 lists the duration values for their component parts.

Note that when pre-aspiration is included in the duration of the consonant, the C/V ratio is 2.1 in both stimuli. This value is well above the ratio of 1.0 required for native Italians to perceive a geminate. A 10-step continuum to singleton *fato*

Table 2: Duration values for the stimulus at the geminate endpoint of the plain and pre-aspirated *fatto-fato* continua. The column /tː/ indicates the duration of the consonant portion including pre-aspiration.

	/a/	Pre-asp.	Closure	VOT	/tː/	C/V
[fatːo]	78	0	156	8	164	2.1
[faʰtːo]	78	48	108	8	164	2.1

6 Twenty-eight participants took part in the production experiment. Seven participants did not complete the perception experiment and were therefore excluded. One participant (later discovered to be bilingual in English) was excluded along with four considerably older participants (between 46 and 61 years of age). Initial comparisons between old ($n = 4$) and young speakers ($n = 16$) showed differences in the perceptual boundary and need to be investigated with a larger sample of older participants.

was then synthesized by manipulating the C/V ratio from 2.1 (= *fatto*) down to 0.6 (= *fato*) in 10 equal steps. To do this, vowel and oral closure durations were manipulated in Praat – not pre-aspiration duration, which was instead kept stable at zero along the plain stop [fatːo] continuum and at 48 ms along the pre-aspirated [faʰtːo] continuum. The 160 stimuli (10 steps × 2 continua × 8 repetitions) were randomized and presented to listeners in a two-alternative forced-choice perception experiment. Listeners completed the experiment on line and were instructed to wear headphones. Their task was to decide whether each stimulus sounded more like *fato* or *fatto* and to click the corresponding button (the buttons showed orthographic words). Participants could listen to the stimuli as many times as they wished.

2.1.3 Data analysis

Listener response data were analysed with the lme4 package in R, following procedures described in Kleber et al. (2011). 50% *fatto-fato* category boundaries were obtained with logistic regression by fitting a generalized linear model to the response data with Listener Response (2 levels 'fatto'/'fato') as the dependent variable, Stimulus Number (10 levels) as a fixed factor, and Listener (15 levels) as a random factor. This fitted sigmoid curves to each listener and each continuum (giving 15 × 2 category boundary values (u), derived from the listener-specific intercept (m) and slope (k) values by calculating $u = -k/m$). The influence of pre-aspiration on the *fatto-fato* category boundary u was tested with a mixed model within R package lme4 (Bates 2010) with u as the dependent variable, Continuum (2 levels 'plain'/'pre-aspirated') as a fixed factor, and Listener as a random factor. The p value was obtained by comparing the full model to a null model.

2.2 Results

Figure 2 shows listener responses to the two *fato-fatto* continua (combined for all listeners). The proportion of 'fatto' responses increases as the C/V ratio increases. For the plain continuum (labelled 'fa[tː]o'), the 50% category boundary for geminate *fatto* is just under the threshold value of 1.0 suggested by Pickett et al. (1999). This means that even when the consonant was slightly shorter than the preceding vowel, listeners in this study still perceived it as a geminate. Responses to the pre-aspirated continuum ('[fa[ht]o') show a rightwards shift.

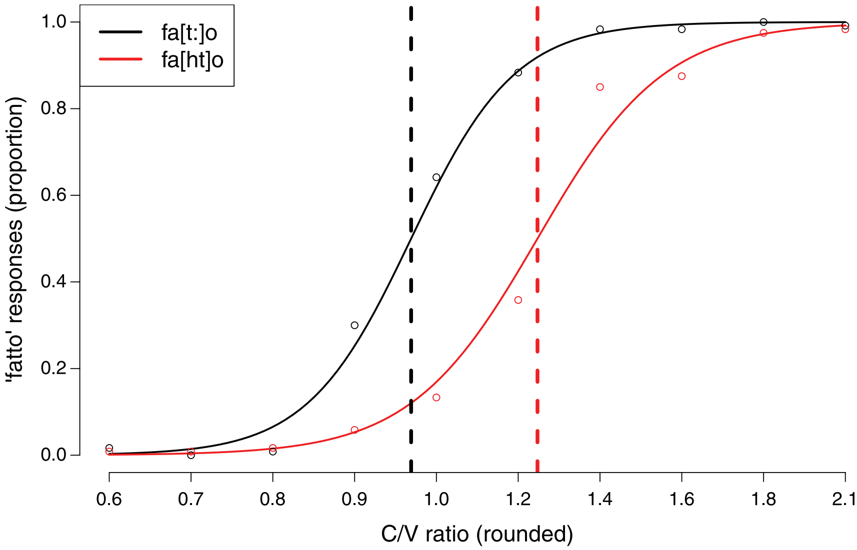

Fig. 2: Proportion of geminate responses for the four *fato-fatto* continua (averaged on 15 listeners). The vertical dashed lines indicate the average 50% perceptual boundaries. The x-axis shows the 10 stimuli labelled according to C/V ratio value (Stimulus 1 = 0.6, Stimulus 10 = 2.1).

The fact that there is a difference between the two continua means that pre-aspiration affected listener categorization of phonological consonant length. The fact that the pre-aspirated continuum shifts to the right means that pre-aspirated stops are less likely to be perceived as geminates than their plain counterparts. In other words, replacing part of the oral closure with pre-aspiration elicited more *fato* responses. This result clearly shows that listeners did not parse pre-aspiration with the oral closure in categorizing the stimuli, because had they done so we would have found no difference between the two continua. The mixed model conducted with the 50%-perceptual category boundary as the dependent variable, Continuum as a fixed factor, and Participant as a random factor confirmed that the perception of consonant length is significantly affected by pre-aspiration ($\chi^2[1] = 31.56$, $p < 0.001$). Inspection of individual listener data also showed a later 50% category boundary – if only slightly – in response to the pre-aspirated continuum in all cases. Most listeners showed a difference of between 1 and 2.5 continuum steps, but for two listeners the difference was less than 1 continuum step, i.e., the pre-aspirated continuum showed only a very slight rightwards shift. Such small differences between the category boundaries for the plain and pre-aspirated continua suggest that these two listeners tended to parse pre-aspiration with consonant closure.

Table 3: Segment durations (vowel "V", pre-aspiration "Pre", and closure+VOT "C", in ms) for each of the ten stimuli along the pre-aspirated *fato-fatto* continuum. $(C_{+pre})/V$ and $C/(V_{+pre})$ show C/V ratios with pre-aspiration duration included in the consonant and in the vowel, respectively. The percentage of geminate *fatto* responses to each stimulus (corresponding to the red curve in Figure 2) is listed in the right-hand column.

Stim. no.	V	Pre	C	$(C_{+pre})/V$	$C/(V_{+pre})$	% *fatto*
1	151	48	92	0.6	0.5	0.8
2	143	48	100	0.7	0.5	0.8
3	135	48	108	0.8	0.6	1.7
4	127	48	116	0.9	0.7	5.8
5	119	48	124	**1.0**	0.7	**13.3**
6	111	48	132	1.2	0.8	35.8
7	103	48	140	1.7	0.9	85.0
8	95	48	148	1.6	**1.0**	**87.5**
9	87	48	156	1.8	1.2	97.5
10	79	48	164	2.1	1.3	98.3

Nonetheless, most listeners clearly did *not* parse pre-aspiration with the consonant closure for the purposes of computing a C/V ratio in order to categorize stimuli as *fato* or *fatto*. The question now arises as to whether they parsed it with the vowel or independently. While our experiment was not designed to test this issue – concentrating instead on the impact of pre-aspiration on consonant length – closer inspection of the response data can shed some light on whether listeners parsed pre-aspiration with the vowel or independently. More specifically, Table 3 lists the C/V ratio $(C_{+pre})/V$ with pre-aspiration duration included in the consonant, as has been the case up to this point, and $C/(V_{+pre})$ with pre-aspiration included in the vowel.

The proportion of geminate responses in the right-most column should track C/V ratio values because the C/V ratio is a reliable cue to phonological length for Italian listeners. We just saw that participants did not parse pre-aspiration with oral closure, $(C_{+pre})/V$, for the purposes of computing a C/V ratio in this perception task because their responses differed between the two continua in Figure 2. This is also illustrated in Table 3, because for the $(C_{+pre})/V$ column, the geminate threshold lies at stimulus 5, but listeners require a C/V ratio of at least 1.7 to respond with *fatto*. The next question is whether listeners parsed pre-aspiration with the vowel for the purposes of computing a C/V ratio and making a decision about phonological length. C/V ratio values with pre-aspiration included together with that of the vowel are listed in the $C/(V_{+pre})$ column. We can see that in this case the geminate threshold lies at stimulus 8 (also in bold). In the right-most column, there is a considerable jump in the percentage of geminate responses from stimulus 6 (35.8%) to stimulus 7 (87.5%), yet both of these stimuli are below the perceptual threshold of 1.0, if pre-aspiration is parsed to the vowel. Therefore, this measure corresponds no better to listener responses than when pre-aspiration was parsed with the consonant closure, $(C_{+pre})/V$. Thus, based on these data, it appears that our Italian listeners parsed pre-aspiration independently from the vowel and independently from the oral closure when calculating the C/V ratio for the purposes of making a decision about phonological length.

3 Production experiment

3.1 Method

The production experiment was designed to assess whether pre-aspiration involves concomitant oral closure shortening in the production of geminate stops. To do so, multiple repetitions of words containing voiceless geminates were recorded and divided according to whether they were produced with or without intervening pre-aspiration.[7] The hypothesis is that oral closure should be shorter for the pre-aspirated tokens.

3.1.1 Participants

Participants were the same as those who participated in the perception experiment.

3.1.2 Data collection

Recordings were made by the first author in a private living room in Prato using the SpeechRecorder software (Draxler and Jänsch 2004) on a MacBook Pro and a headset microphone. Participants were given no explicit instruction as to whether they should speak with the local accent or in standard Italian. This is because the two are not necessarily separable in practice for native Tuscan speakers

7 Geminates are not compared to singleton stops because they are not directly comparable on any acoustic measure: intervocalic singletons can be realised as voiceless stops or voiceless oral or glottal fricatives or completely elided, e.g., /t/ in *prato* 'field' > [t]~[θ]~[h]~[ø] under the *Gorgia toscana* (cf. Section 2.1.1).

(cf. Bertinetto and Loporcaro 2005: 144) and the words were to be read in as natural a manner as possible. This is important given pre-aspiration, as a novel synchronic variant, might not occur at all in more formal settings. The data involve a subset collected towards a larger investigation for which participants read a list of real words containing post-tonic /p t k pː tː kː/ in various contexts. The subset analysed here involves words containing geminate /tː/ (*ditta, matti, fatto, gatta* 'fingers, crazy, done, female cat'), facilitating comparison with the perception data which also involved /tː/. With repetitions, there were 60 /tː/ tokens per speaker. Two were discarded due to background noise, giving 898 tokens in total.

3.1.3 Labelling

Word tokens were labelled semi-automatically with MAuS (Schiel et al. 2011), and the acoustic boundaries were then adjusted manually using Praat (Boersma and Weenink 2012). For each token the duration of the following portions was measured:
- Vowel
- Pre-aspiration
- Closure
- VOT (= burst and post-aspiration)

The presence of pre-aspiration in individual tokens was decided manually based on auditory as well as acoustic cues, taking the offset of modal voicing in the vowel as the primary acoustic cue (i.e., breathy voicing, where present, was included in the pre-aspiration portion). This is consistent with previous work on Italian (e.g., Stevens 2011) and follows standard procedures for labelling pre-aspiration (e.g., Ní Chasaide 1985; Helgason 2002; van Dommelen and Ringen 2007; Helgason and Ringen 2008). Where no pre-aspiration intervened in the vowel-voiceless stop transition, pre-aspiration duration was recorded as zero. These duration data were imported into Emu/R (Harrington 2010) for analysis.

3.1.4 Analysis

The segmental affiliation of pre-aspiration in the production of vowel + /tː/ sequences is tested with primary reference to the C/V ratio. This measure was defined following Pickett et al. (1999) except that voice onset time was included with closure duration in the present study, i.e., C/V = (closure duration + VOT)/

vowel duration. Pickett et al. (1999) did not report pre-aspiration in their study and defining the C/V ratio in this way means that the duration of any intervening pre-aspiration in the present data is removed from the equation.

(2) Word	Produced as	C/V ratio calculated as
fatto	[fatːo]	[t] dur./[a] dur.
fatto	[faʰtːo]	[t] dur./[a] dur.

Excluding pre-aspiration from the C/V ratio allows us to investigate pre-aspiration in production. Recall the available acoustic evidence (cf. Table 1) showing that pre-aspiration appears to involve concomitant oral closure shortening. If this is also true for more controlled comparisons, then C/V ratios should be lower for the tokens realised with a portion of pre-aspiration [faʰtːo] than for their plain counterparts [fatːo]. This is because excluding pre-aspiration from the C/V for [faʰtːo] tokens implies subtracting pre-aspiration from the numerator. If, on the other hand, speakers time pre-aspiration with the vowel, then C/V ratios should be *higher* for the tokens produced with pre-aspiration. This is because subtracting pre-aspiration duration from the denominator (V) would increase the C/V ratio for the pre-aspirated tokens. Finally, pre-aspiration might be independent from quantity distinctions (cf. Section 1.5). If this is so, then leaving pre-aspiration duration out of the C/V ratio equation should make no difference, and values should be similar for the pre-aspirated and plain tokens. Table 4 summarizes these three hypotheses and what they mean in terms of the temporal affiliation of pre-aspiration in the production of vowel + /pː tː kː/ sequences.

A mixed model within R package lme4 (Bates 2010) was applied to these data with one of either Vowel Duration, Closure + VOT Duration, or C/V Ratio as the dependent variable, Pre-aspiration as a fixed factor (2 levels: pre-aspirated or plain), and Speaker (15 levels) and Word (4 levels) as random factors. In each case we obtained *p* values by running the model a second time without pre-aspiration as a fixed factor and then comparing the two models with a Chi-squared test.

Table 4: Three potential scenarios regarding the C/V ratio for pre-aspirated ($C/V_{pre\text{-}asp}$) and plain (C/V_{plain}) tokens. The right-hand column lists the implications.

	C/V ratio	Implication
A	$C/V_{pre\text{-}asp} < C/V_{plain}$	Speakers co-ordinate pre-aspiration temporally with oral closure
B	$C/V_{pre\text{-}asp} > C/V_{plain}$	Speakers co-ordinate pre-aspiration temporally with the vowel
C	$C/V_{pre\text{-}asp} \approx C/V_{plain}$	Pre-aspiration is timed independently in production

3.2 Results

Pre-aspiration occurred in 155 of the 898 (17.3%) tokens. Individuals varied to the extent that they pre-aspirated in the production of /tː/: speakers ranged from producing no pre-aspiration at all (two speakers) to producing 37% of their /tː/ tokens with pre-aspiration. In Figure 3 the number of pre-aspirated tokens is plotted against mean pre-aspiration (one data point per participant).

The frequency and duration of pre-aspiration are positively correlated: not all speakers pre-aspirated, and those that did so more often showed longer pre-aspiration portions. This correlation between the frequency and duration of pre-aspiration was accounted for by including speaker as a random factor in the mixed model. Moreover, only data for participants whose tokens included both pre-aspirated and plain stops were plotted. This was done to avoid biasing comparisons between plain and pre-aspirated tokens to those individuals who pre-aspirated more frequently. Thus, the number of data points (each representing mean duration for one speaker) varies across comparisons and is listed at the top of each cell in Figures 4–6.

Vowel duration is shorter with pre-aspiration, and the mixed model with vowel duration as the dependent variable confirms that pre-aspiration has a sig-

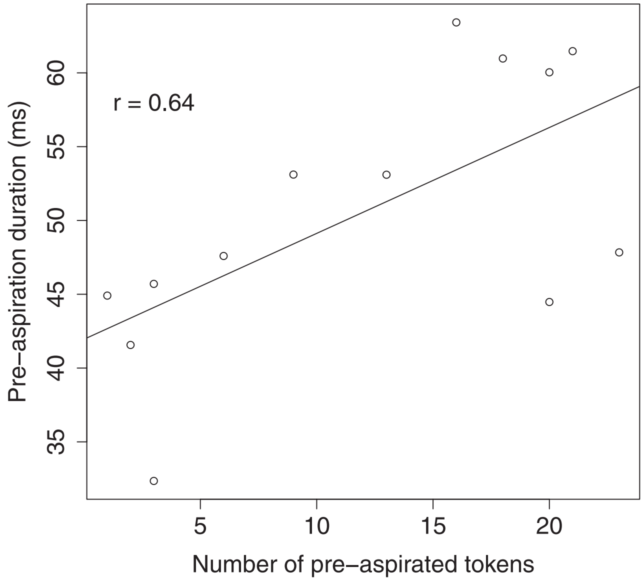

Fig. 3: For each speaker participant, the number of pre-aspirated tokens plotted against the mean duration of pre-aspiration.

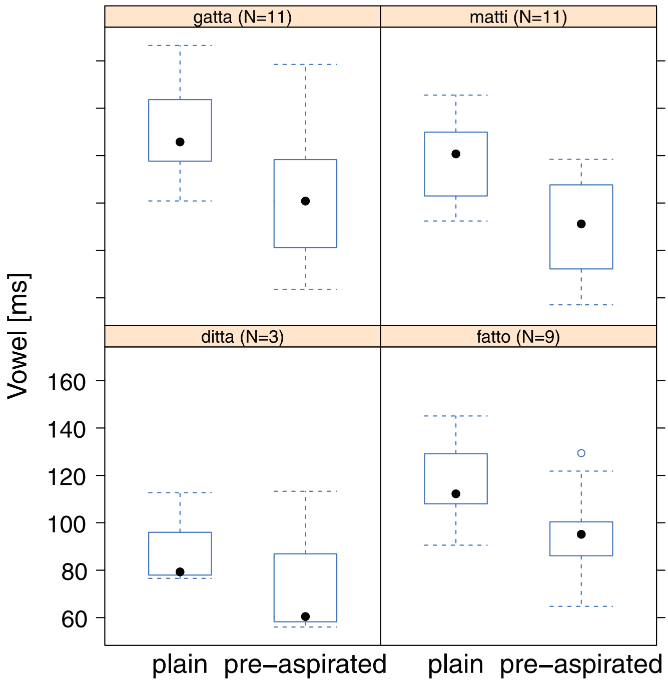

Fig. 4: Boxplots for mean vowel duration (in ms) preceding plain and pre-aspirated /tː/, separated according to word identity. N = number of speakers who contributed data points.

nificant effect on vowel duration ($\chi^2[1] = 108.9$, $p < 0.001$). This pattern conflicts with results based on less controlled comparisons for pre-aspiration in Italian (cf. §1.3) which instead suggest that vowel duration remains relatively stable whether or not pre-aspiration intervenes in a vowel + /pː tː kː/ sequence. Figure 5 shows that the mean duration of pre-aspiration varies according to vowel type: it is shorter in *ditta* than following low /a/ in the other three words.

Pre-aspiration is also significantly shorter after high vowels in other languages (e.g., Nance and Stuart-Smith 2013: 137 on Scottish Gaelic). The effect of vowel type on pre-aspiration duration in Italian was tested with a mixed model on the subset of 155 pre-aspirated tokens, with Pre-aspiration Duration as the dependent variable, Vowel Type (2 levels) as the fixed factor, and Speaker as a random factor. This model confirmed that vowel type has a significant influence on pre-aspiration duration ($\chi^2[1] = 20.731$, $p < 0.001$). The pattern in Figure 5 also reflects the correlation between frequency and duration of pre-aspiration in these data: only three participants produced *ditta* with pre-aspiration, whereas nearly all participants produced pre-aspiration at least once in the /a/ words. In

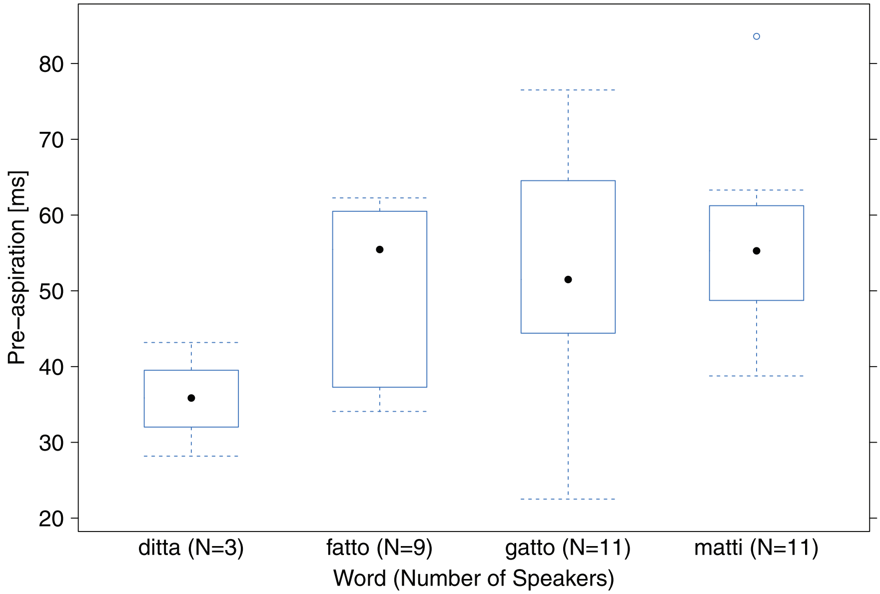

Fig. 5: Pre-aspiration duration (in ms) by word. N = number of speakers who contributed data points.

contrast to these patterns for vowel duration, pre-aspiration had little impact on the duration of the voiceless stop /tː/: mixed models showed no influence of pre-aspiration on closure duration or on VOT duration. The summed duration of Closure + VOT also remained stable whether or not pre-aspiration occurred, and a mixed model with Closure + VOT as the dependent variable confirmed that pre-aspiration had no effect on the duration of the stop ($\chi^2[1] = 0.0188$, $p = 0.8911$). Mean C/V values were consistently higher for the pre-aspirated tokens, as seen in Figure 6, and a mixed model applied with C/V Ratio as the dependent variable confirmed this pattern ($\chi^2[1] = 99.9$, $p < 0.001$).

Recall from Section 3.1.4 that pre-aspiration was not included in the C/V ratio calculations, which led to three scenarios about the temporal affiliation of pre-aspiration in the production of vowel + /pː tː kː/ sequences. The pattern in Figure 6, with higher C/V ratios for the pre-aspirated tokens, can be interpreted as follows: adding pre-aspiration to the denominator would decrease the C/V ratio, bringing pre-aspirated stops more into line with plain stops on this measure. This favours the interpretation that pre-aspiration is temporally co-ordinated with the vowel. This pattern was upheld within individual speakers, i.e., most showed higher mean C/V ratios for their tokens produced with pre-aspiration. However, differences between plain and pre-aspirated tokens

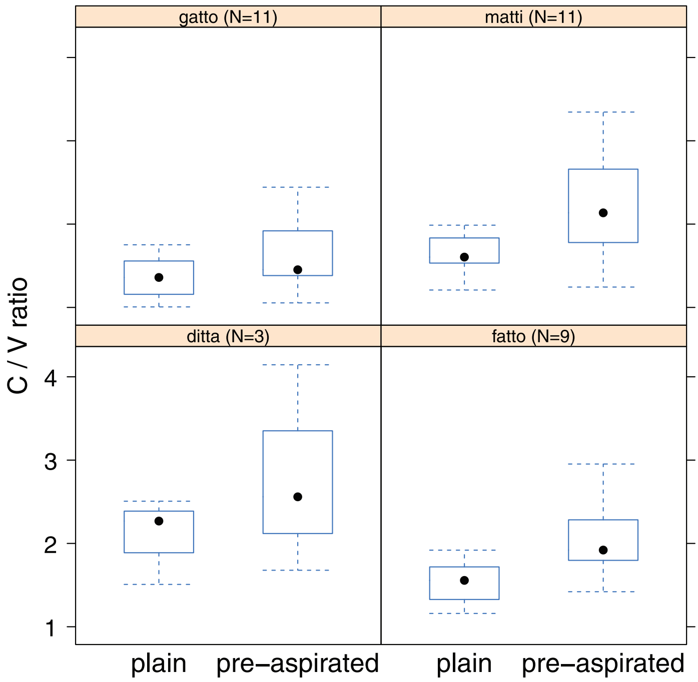

Fig. 6: C/V ratio for plain and pre-aspirated tokens. N = number of speakers who contributed data points: within each cell only data points for speakers who showed both pre-aspirated and plain tokens were plotted. Mean C/V ratio values for plain and pre-aspirated stops are well above Pickett et al.'s (1999) suggested geminate threshold of 1.0 and can be interpreted as long on this measure.

were often very small, with several speakers showing no clearly visible difference between plain and pre-aspirated tokens on this measure. Nonetheless, no speaker showed lower mean C/V values for their pre-aspirated tokens. Therefore, and in contrast with the predictions of this study, there is no evidence that pre-aspiration involved concomitant closure shortening for any of the speakers.

4 Discussion

4.1 Pre-aspiration and quantity in perception

This study tested Italian listener perception of pre-aspirated stops, which have been recently documented to occur in the ambient language. Results showed that when part of the oral closure for a voiceless geminate stop /tː/ is replaced with pre-aspiration, Italian listeners are more likely to perceive a short stop /t/. For all but two of our listeners, only long closure was sufficient to elicit geminate responses. This result supports the perceptual part of our hypothesis about pre-aspiration in Italian, i.e., that listeners would tend not to parse pre-aspiration with oral closure (as described at (1c)). It is also in line with other experimental studies on Italian (Esposito and Di Benedetto 1999; Pickett et al. 1999), which report that native listeners attend primarily to oral constriction duration in distinguishing consonant length (but which have not addressed aspiration). The prediction made in earlier studies on pre-aspiration in Italian (Stevens 2011, 2012), namely that Italian listeners might show a trading relation between pre-aspiration + oral closure in making perceptual quantity judgements was not supported by the experimental data reported here.

Some variability across listener responses was expected given that experimental research shows that even stable segmental categories show idiosyncratic mappings with the auditory signal (e.g., Hazan and Rosen 1991; Beddor 2009; Yu 2013; and see also Section 4.4). While most listeners parsed pre-aspiration with the vowel, perception data for two listeners were interpreted as their having parsed pre-aspiration with oral closure. These two different parsing strategies for pre-aspiration in the perception data correspond to dialect differences noted earlier for Swedish, where a pre-aspirated V + C sequence can be interpreted as either VCː or VːC. This suggests a link between listener-specific perception of pre-aspiration in the auditory signal and permanent dialect-level categorical differences that is worthy of further investigation.

4.2 Pre-aspiration and quantity in production

The production study confirmed that pre-aspiration is a shared tendency for vowel + /tː/ sequences for some, but not all, Italian speakers. Based on published acoustic data for Italian (cf. Section 1.3) and on the descriptive literature on pre-aspiration (e.g., Thráinsson 1978 on Icelandic; Ó Baoill 1980 on Irish Gaelic; Blevins and Garrett 1993; Ladefoged and Maddieson 1996), we expected pre-aspirated tokens to show concomitant oral closure shortening for /pː tː kː/. The hypothesis described in (1c) was that Italian listeners would attend to (shorter) oral closure and thus interpret pre-aspirated pronunciations as short /p t k/. However, the production data showed that oral closure duration remained stable across the plain and pre-aspirated tokens. In other words, our speaker subjects did not produce the kind of synchronic variation for /pː tː kː/ that forms the first stage of our proposed model of de-gemination via pre-aspiration. Instead, whether or not /pː tː kː/ were produced with pre-aspiration, oral closure duration was always sufficiently long to signal a geminate /tː/. Because there was no evidence in the production data to suggest that speakers produce shorter oral closures with pre-aspiration, our hypothesis that pre-aspiration could lead to de-gemination of /pː tː kː/ in Italian must be rejected.

The discrepancy between the present production data and earlier results for Italian, which suggested that concomitant oral closure shortening occurs with pre-aspiration, can be explained by methodological differences. Earlier published data were drawn from spontaneous speech (Stevens 2012) or from word-list recordings included within the CLIPS corpus (Stevens 2010, 2011). The CLIPS data, in particular, involved a large number of speakers but only one token per word per speaker. The present data involved a higher number of repetitions per speaker, and stop place and surrounding context were controlled.

The fact that pre-aspiration does not involve concomitant closure shortening in our Italian production data has potential implications for theories about the origins of pre-aspiration in geminate /pː tː kː/. Some sources (cf. Section 1.2) speculate that pre-aspiration originates as the unintended consequence of shorter oral closures in the production of /pː tː kː/; this pre-aspiration may then be noticed by listeners and chosen for sound change as in (1b). Ní Chasaide's (1985: 294) diachronic model is instead based on the idea that devoicing of /bː dː gː/ might be the reason why speakers come to produce /pː tː kː/ with pre-aspiration, in order to keep the two geminate stop series distinct in production and perception. We did not collect production data for /bː dː gː/ in Italian, so it is not possible to test whether they may be undergoing devoicing and whether this may help to explain pre-aspiration in the voiceless series. However, Ní Chasaide makes the suggestion that pre-aspiration might initially serve to increase the overall duration of syllables containing /pː tː kː/ and that closure shortening may be a subsequent development. The production data for Italian reported here support this suggestion inasmuch as optional pre-aspiration does not involve concomitant closure shortening for /pː tː kː/. Moreover, the duration of the entire vowel + /tː/ sequence is about 30 ms longer with pre-aspiration, that is, pre-aspiration does increase the duration of the V + Cː sequence. A mixed model as outlined in Section 3.1.4 but with V + /tː/ Duration as the dependent variable, +/− Pre-aspiration as the fixed factor, and Word and Speaker as random factors confirmed this pattern ($\chi^2[1] = 31.245$, $p < 0.001$). This result lends support to Ní Chasaide's suggestion that pre-aspiration precedes closure shortening rather than the other way around.

4.3 Pre-aspiration and sound change

Van Dommelen (1999: 2040) outlines a two-stage model for the development of pre-aspiration in a language. He observes that some Norwegian dialects show breathy voice at the vowel offset preceding voiceless stops, but only the "pre-aspirating dialects" show the second stage, which involves "longer and mainly voiceless glottal frication". Helgason (2004: 3) also proposes a two-stage model for pre-aspiration in Swedish, including the idea that pre-aspiration might originate in accidentally mistimed gestures and then develop into a planned feature for word-medial fortis stops /pː tː kː/. Helgason describes how pre-aspiration duration varied consistently according to the phonemic length of the preceding vowel for two speakers, suggesting it is a planned event, whereas a third speaker showed less frequent and much shorter pre-aspiration durations, which Helgason attributes to incidental differences in inter-gestural timing. Helgason relates these two timing patterns to Lindblom's H & H theory (e.g., Lindblom et al. 1995), describing the former, planned event as a hyper-variant and the latter as a hypo-realisation. Helgason suggests that only the hyper-variant would drive the sound change leading to obligatory pre-aspiration.

Within this kind of two-stage model, the Italian production data suggest that pre-aspiration is best interpreted as a hyper-variant. This is because pre-aspiration involves early glottal abduction *in addition* to a long oral closure portion in the production of /tː/, making the pre-aspirated tokens more 'geminate-like' than their plain counterparts. On the other hand, the perception data suggest that the Italian participants had difficulty detecting pre-aspiration (even though pre-aspiration was completely voiceless in the stimuli), and did not parse it with oral closure. This suggests that Italian listeners do not interpret pre-aspirated stops as a hyper-variant of geminate /tː/, of the kind that would precipitate sound change to obligatory pre-aspiration within Helgason's perceptually driven model.

The tendency for Italian listeners not to parse pre-aspiration with the following consonant closure can be attributed to lack of experience with contrastive /h/, in line with Mielke's (2003) account of the influence of native phonology on the ability to detect glottal frication (as noted earlier at Section 1.5). Taken together with Mielke's results, the present perception data for Italian suggest that the typological rarity of pre-aspiration might be explained not only by the auditory ambiguity of pre-aspiration itself (Silverman 2003), but also by structural differences between languages. That is, pre-aspiration might be a synchronic production variant in more languages than is currently known to be the case, but native listeners need experience with contrastive /h/ in order for pre-aspiration to develop into a contrastive feature.

4.4 Individual differences in the production and perception of pre-aspiration

The production data showed that not all participants produced pre-aspiration, and the perception data showed that perceptual parsing strategies for pre-aspiration differed amongst participants. Because the same participants took part in both experiments, it is worth exploring whether these differences are due to idiosyncratic – but consistent – parsing strategies or whether the parsing of pre-aspiration in production and perception varies within the same individual. Grosvald and Corina (2012: 96) suggest that (only) a small proportion of the community might be especially sensitive to novel variants and likely to reproduce them, and that it is these individuals who drive sound change. With this in mind, we are interested in whether the two individuals who parsed pre-aspiration with oral closure in perception might match this pattern in their own productions (i.e., producing shorter closure durations with pre-aspiration). This would suggest that they are using pre-aspiration consistently in the perception and production of quantity contrasts in Italian and that they might be considered innovators in a sound change involving the replacement of part of the oral closure for /tː/ with pre-aspiration.

With Grosvald and Corina's (2012) study as a model, the production and perception data were plotted against one another in Figure 7. The data are divided into four quadrants, intended to separate participants according to the way they perceived and produced pre-aspirated stops. To reduce the perception data to one value, each listener's 50% category boundary value for the plain continuum was subtracted from that in response to the pre-aspirated continuum (cf. Figure 2 for the category boundaries). Values close to zero indicate that replacing oral closure with pre-aspiration made little difference to the *fatto-fato* category boundary for that listener; higher values mean more *fato* responses for the pre-aspirated continuum (i.e., pre-aspiration was not parsed with oral closure). Thus the perception data for individual listeners are plotted on a scale from "pre-aspiration + closure" to "closure", which corresponds to whether that listener parsed pre-aspiration with oral closure, or not, in categorizing stimuli as *fatto* or *fato*. The production data were reduced to one value per speaker by subtracting each speaker's mean C/V for plain tokens from their mean C/V for pre-aspirated tokens (here excluding pre-aspiration from the C/V calculations following the methods of the production experiment; data for speakers who did not pre-aspirate were excluded from Figure 7). C/V_{pre}-C/V_{plain} values close to zero mean there was no difference between the two pronunciations. That is, a speaker with a value of zero on the y-axis co-ordinated pre-aspiration independently of the vowel and independently of the oral closure in production. Positive values on

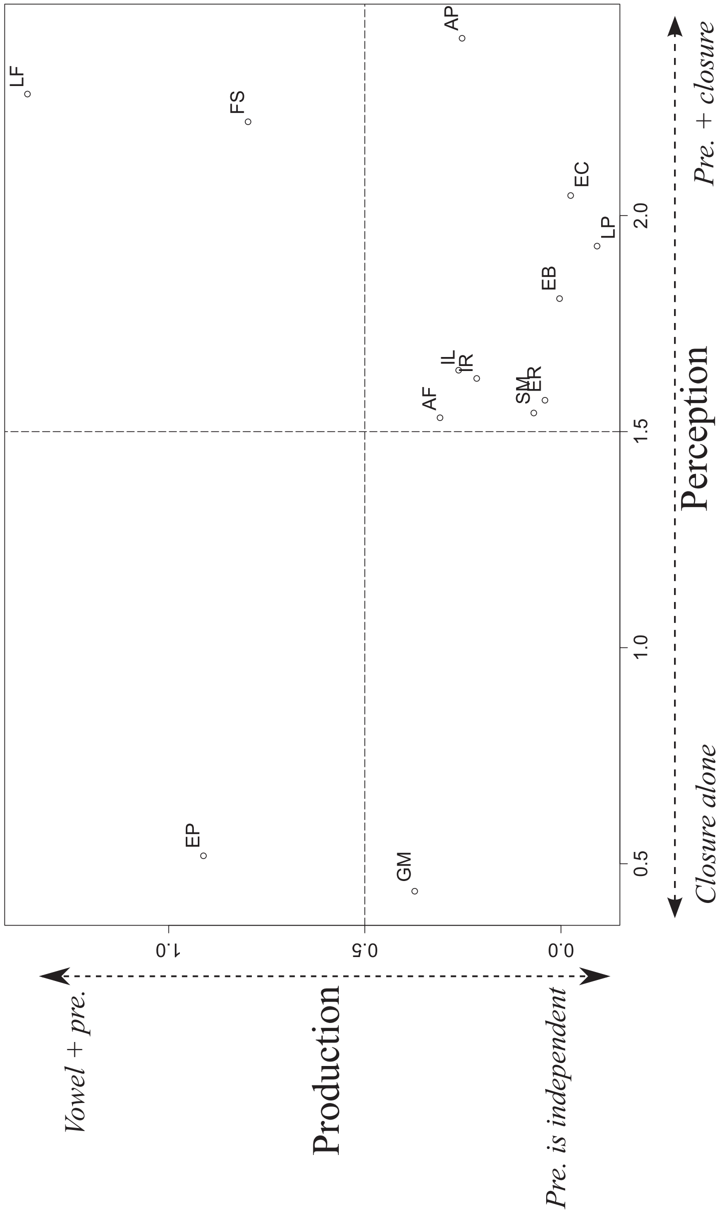

Fig. 7: Participant-specific patterns for pre-aspiration in production (y-axis) and in perception (x-axis). Both axes show differences between pre-aspirated and plain stop tokens in terms of C/V. See text.

the y-axis reflect a higher C/V_{pre}, that is, for that particular speaker, pre-aspiration is temporally affiliated with the vowel. Negative values for a particular speaker would mean that pre-aspiration involves concomitant shortening of oral closure, but as we saw in Section 3, this was not the case for any of the speakers in the production study.

Overall, pre-aspiration appears to be more stable in perception – with most participants in the right-hand quadrants – than in production, for which participants are spread relatively evenly over the y-axis. The relative stability of pre-aspiration in perception relative to production appears to be a feature of pre-aspiration itself, rather than being due to any potential differences between experimental groups (given that the same participants took part in both experiments). There is evidence that pre-aspiration is less stable in production than in perception in Scandinavian varieties where it is an optional feature. For example not all speakers of Central Standard Swedish tend to pre-aspirate, but in perception Helgason (2004) found that all 16 listeners used pre-aspiration to distinguish, e.g., *ratta* 'steer' and *radda* 'a lot of'.

Most listeners (all except GM and EP) are on the right of the threshold line at 1.5 in perception. That is, their category boundary for the pre-aspirated continuum shifted to the right (more *fato* responses) by at least 1.5 continuum steps. In terms of their own /tː/ productions, most of these participants fell in the bottom quadrant. This means they produced pre-aspiration independently of vowel and oral closure. The remaining two participants in Figure 7 (GM and EP) perceived little difference between plain and pre-aspirated stimuli, i.e., they parsed pre-aspiration with closure in perception. GM and EP did not match this pattern in production, however, for which y-axis values should have been below zero. Therefore, Figure 7 suggests that GM and EP show a trading relation between pre-aspiration and oral closure in perception but not in their own production. As such these two participants are much like all the others in showing a mismatch between the production and perception of pre-aspiration. Indeed Figure 7 shows that none of the participants in this study uses pre-aspiration systematically in the production and perception of quantity.

5 Conclusions

The hypothesis tested in this study was that synchronic variability involving the production of /pː tː kː/ with pre-aspiration and somewhat shorter oral closure would lead to de-gemination of /pː tː kː/ in Italian. Within the terms of Ohala's (1981, 1993) perceptually driven model of sound change, the proposal was that listeners would fail to parse pre-aspiration with the oral closure, thus mistaking pre-aspirated /pː tː kː/ for short /p t k/. The perception data supported this hypothesis: replacing part of the closure duration with pre-aspiration elicits more singleton responses from Italian listeners. However, the production data showed that pre-aspiration does not involve concomitant closure shortening. In other words, Italian speakers are not producing the synchronic variation for /pː tː kː/ that would cause native listeners to perceive them as short (in our model). This means that the hypothesis that Italian might be undergoing listener-driven degemination must be rejected. The production data lend support to Ní Chasaide's (1985: 297) suggestion that the link between pre-aspiration and quantity in historical sound change is an indirect one. In particular, there was no evidence in this study of a trading relationship between oral closure and pre-aspiration in the production or perception of geminate voiceless stops.

Acknowledgments: We are very grateful for help with this research from the editor Jonathan Barnes, as well as from three anonymous reviewers and from discussions with Jonathan Harrington and Florian Schiel. Matthew Absalom helped recruit participants in Prato, and Florian Schiel created the web interface for the perception experiment. Earlier versions of this research were presented at the Workshop on Sound Change Actuation in Chicago in April 2013 and at the Empirical Approaches at the Boundary of Phonetics and Phonology seminar in Venice in May 2013. This research was funded by ERC grant 295573 to Jonathan Harrington.

References

Bates, Douglas M. 2010. *Mixed-effects modelling with R*. http://lme4.r-forge.r-project.org/lMMwR/lrgprt.pdf

Beckman, Mary, Kenneth De Jong, Sun-Ah Jun, & S.-H. Lee. 1992. The interaction of coarticulation and prosody in sound change. *Language and Speech* 35. 45–58.

Beddor, Patrice Speeter. 2009. A coarticulatory path to sound change. *Language* 85(4). 407–428.

Beddor, Patrice Speeter. 2012. Perception grammars and sound change. In María-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change: perception, production, and social factors*, 37–55. Amsterdam/Philadelphia: John Benjamins.

Bertinetto, Pier Marco, & Michele Loporcaro. 2005. The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association* 35(2). 131–151.

Bladon, Anthony. 1986. Phonetics for hearers. In Graham MacGregor (ed.), *Language for hearers*, 1–24. Oxford: Pergamon.

Blevins, Juliette, & Andrew Garrett. 1993. The evolution of Ponapeic nasal substitution. *Oceanic Linguistics* 32(2). 199–236.

Boersma, Paul, & David Weenink. 2012. *Praat: doing phonetics by computer* [computer program]. Version 5.3.27. http://www.praat.org

CLIPS corpus (Corpora e Lessici di Italiano Parlato e Scritto). http://www.clips.unina.it

Dalcher, Christina Villafana. 2008. Consonant weakening in Florentine Italian: a cross-disciplinary approach to gradient and variable sound change. *Language Variation and Change* 20. 275–316.

Draxler, Christoph, & Klaus Jänsch. 2004. SpeechRecorder – A universal platform independent multichannel audio recording software. *Proceedings of the Fourth International Conference on Language Resources and Evaluation,* Lisbon, Portugal.

Esposito, Anna, & Maria Di Benedetto. 1999. Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America* 106(4). 2051–2062.

Foster, Mary LeCron. 1969. *The Tarascan language* (University of California Publications Linguistics 56). Berkeley & Los Angeles: University of California Press.

Foulkes, Paul, Gerry Docherty, & Dominic Watt. 1999. Tracking the emergence of sociophonetic variation. *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco. 1625–1628.

Giannelli, Luciano, & Thomas Cravens. 1997. Consonantal weakening. In Martin Maiden & Mair Parry (eds.), *The dialects of Italy*, 32–40. London/New York: Routledge.

Grosvald, Michael, & David Corina. 2012. The production and perception of sub-phonemic vowel contrasts and the role of the listener in sound change. In María-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change: Perception, production, and social factors*, 77–100. Amsterdam/Philadelphia: John Benjamins.

Hansson, Gunnar Ólafur. 2008. Diachronic explanations of sound patterns. *Language & Linguistics Compass* 2(5). 859–893.

Harrington, Jonathan. 2010. *The phonetic analysis of speech corpora*. Oxford: Wiley-Blackwell.

Harrington, Jonathan. 2012. The relationship between synchronic variation and diachronic change. In Abigail Cohn, Cécile Fougeron, & Marie Huffman (eds.), *Handbook of Laboratory Phonology*, 321–332. Oxford: Oxford University Press.

Hawkins, Sarah. 2010. Phonological features, auditory objects, and illusions. *Journal of Phonetics* 38. 60–89.

Hazan, Valerie, & Stuart Rosen. 1991. Individual variability in the perception of cues to place contrasts in initial stops. *Perception and Psychophysics* 49(2). 187–200.

Helgason, Pétur. 2002. Preaspiration in the Nordic Languages: Synchronic and diachronic aspects. Stockholm: Stockholm University Ph.D. dissertation.

Helgason, Pétur. 2004. The perception of medial stop contrasts in Central Standard Swedish: a pilot study. In Peter Branderud & Hartmut Traunmüller (eds.), *Proceedings of FONETIK 2004*, Stockholm University Department of Linguistics. 92–95.

Helgason, Pétur, & Catherine Ringen. 2008. Voicing and aspiration in Swedish stops. *Journal of Phonetics* 36. 607–628.

Hualde, José I., & Mariana Nadeu. 2011. Lenition and phonemic overlap in Rome Italian. *Phonetica* 68. 215–242.

Hura, Susan L., Björn Lindblom, & Randy L. Diehl. 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech* 35(1,2). 59–72.

Janda, Richard D., & Brian D. Joseph. 2003. On language, change and language change. In Richard D. Janda & Brian D. Joseph (eds.), *The Handbook of Historical Linguistics*, 3–180. Malden/Oxford/Melbourne/Berlin: Wiley-Blackwell.

Kleber, Felicitas, Jonathan Harrington, & Ulrich Reubold. 2011. The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech* 55(3). 383–405.
Krämer, Martin. 2009. *The phonology of Italian*. Oxford: Oxford University Press.
Kümmel, Martin J. 2007. *Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen*. Wiesbaden: Reichert.
Ladefoged, Peter, & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
Lindblom, Björn, Susan Guion, Susan Hura, Seung-Jae Moon, & Raquel Willerman. 1995. Is sound change adaptive? *Rivista di Linguistica* 7(1). 5–37.
Loporcaro, Michele. 1996. On the analysis of geminates in Standard Italian and Italian dialects. In Bernhard Hurch & Richard A. Rhodes (eds.), *Natural phonology: The state of the art. Papers from the Bern Workshop on Natural Phonology*, 149–174. Berlin/New York/Amsterdam: Mouton de Gruyter.
Loporcaro, Michele. 2011. Phonological processes. In Martin Maiden, John Charles Smith, & Adam Ledgeway (eds.), *The Cambridge history of the Romance languages: Structures,* Vol. 1, 109–154. Cambridge: Cambridge University Press.
McMurray, Bob, & Ashley Farris-Trimble. 2012. Emergent information-level coupling between perception and production. In Abigail Cohn, Cécile Fougeron, & Marie Huffman (eds.), *The Oxford Handbook of Laboratory Phonology*, 369–395. Oxford: Oxford University Press.
Mele, Biagio. 2009. *Fonetica e fonologica del dialetto di San Giovanni in Fiore*. Zurich: A. Francke.
Mielke, Jeff. 2003. The interplay of speech perception and phonology: experimental evidence from Turkish. *Phonetica* 60. 208–229.
Nance, Claire, & Jane Stuart-Smith. 2013. Preaspiration and postaspiration in Scottish Gaelic stop consonants. *Journal of the International Phonetic Association* 43(2). 129–152.
Ní Chasaide, Ailbhe. 1985. Preaspiration in phonological stop contrasts: an instrumental phonetic study. University of Wales Ph.D. thesis.
Ní Chasaide, Ailbhe, & Christer Gobl. 1993. Contextual variation of the vowel source as a function of adjacent consonants. *Language and Speech* 36(2, 3). 303–330.
Ó Baoill, Donall P. 1980. Preaspiration, epenthesis and vowel lengthening – interrelated and of similar origin? *Celtica* 13. 79–108.
Ohala, John J. 1981. The listener as a source of sound change. In Carrie S. Masek, Roberta A. Hendrick, & Mary Frances Miller (eds.), *Papers from the Parasession on Language and Behavior*, Chicago, Chicago Linguistics Society. 178–203.
Ohala, John. 1993. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, 237–278. London: Longman.
Ohala, John J., & Maria Grazia Busà. 1994. Nasal loss before voiceless fricatives: a perceptually-based sound change. *Rivista di Linguistica* 7. 125–144.
Payne, Elinor. 2005. Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association* 35(2). 153–181.
Pickett, E. R., Sheila E. Blumstein, & Martha W. Burton. 1999. Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica* 56. 135–157.
Pind, Jörgen. 1996. Rate-dependent perception of aspiration and pre-aspiration in Icelandic. *The Quarterly Journal of Experimental Psychology* 49A(3). 745–764.

Rohlfs, Gerhard. 1966. *Grammatica storica della lingua italiana e dei suoi dialetti: fonetica*. Turin: Einaudi.
Romero, Joaquín. 1995. An articulatory view of historical /s/-aspiration in Spanish. *Rivista di Linguistica/Haskins Labs* 7(1). 191–208.
Ruch, Hanna, & Jonathan Harrington. 2014. Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics* 45. 12–25.
Savy, Renata, & Francesco Cutugno. 2009. CLIPS diatopic, diamesic and diaphasic variations in spoken Italian. In Michaela Mahlberg, Victorina González-Díaz, & Catherine Smith (eds.), *On-line Proceedings of 5th Corpus Linguistics Conference*, Liverpool, UK. http://ucrel.lancs.ac.uk/publications/cl2009/213_FullPaper.doc.
Schiel, Florian, Christoph Draxler, & Jonathan Harrington. 2011. Phonemic segmentation and labelling using the MAUS technique. New Tools and Methods for Very-Large-Scale Phonetics Research Workshop, University of Pennsylvania.
Schmid, Stephan. 1999. *Fonetica e fonologia dell'italiano.* Turin: Paravia.
Silverman, Daniel. 1997. *Phasing and recoverability.* New York: Garland.
Silverman, Daniel. 2003. On the rarity of preaspirated stops. *Journal of Linguistics* 39(3). 575–598.
Solé, María-Josep. 2007. Controlled and mechanical properites in speech: a review of the literature. In María-Josep Solé, Patrice Speeter Beddor, & Manjari Ohala (eds.), *Experimental approaches to phonology*, 302–321. Oxford: Oxford University Press.
Stevens, Mary. 2010. How widespread is preaspiration in Italy? A preliminary acoustic phonetic overview. In Susanne Schötz & Gilbert Ambrazaitis (eds.), *Working Papers Department of Linguistics and Phonetics Centre for Languages and Literature Lund University*, 97–102. Lund, Sweden.
Stevens, Mary. 2011. Consonant length in Italian: Gemination, degemination and preaspiration. In Scott M. Alvord (ed.), *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, 21–32. Somerville, MA: Cascadilla Proceedings Project.
Stevens, Mary. 2012. *A phonetic investigation into "Raddoppiamento Sintattico" in Sienese Italian speech.* Bern: Peter Lang.
Stevens, Mary, & John Hajek. 2010. Post-aspiration in standard Italian: some first cross-regional acoustic evidence. In *Proceedings of Interspeech 2010*, 1557–1561. Makuhari, Japan: ISCA.
Stevens, Mary, & Jonathan Harrington. 2014. The individual and the actuation of sound change. *Loquens* 1(1), e003. doi: http://dx.doi.org/10.3989/Loquens.2014.003
Thráinsson, Höskuldur. 1978. On the phonology of Icelandic preaspiration. *Nordic Journal of Linguistics* 1(1). 3–54.
Torreira, Francisco. 2012. Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association* 42. 49–63.
van Dommelen, Wim A. 1999. Preaspiration in intervocalic /k/ vs. /g/ in Norwegian. *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco. 2037–2040.
van Dommelen, Wim A., Snefrid Holm, & Jacques Koreman. 2011. Dialectal feature imitation in Norwegian. *Proceedings of the 17th International Congress of Phonetic Sciences,* Hong Kong. 599–602.

van Dommelen, Wim A., & Catherine Ringen. 2007. Intervocalic fortis and lenis stops in a Norwegian dialect. *Norwegian University of Science and Technology Speech, Music and Hearing Quarterly Progress and Status Report* 50(1). 5–8.
Wretling, Pär, E. Strangert, & F. Schaeffler. 2003. Preaspiration as a quantity feature. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona. 2701–2704.
Yu, Alan C. L. 2013. Individual differences in socio-cognitive processing and the actuation of sound change. In Alan C. L. Yu (ed.), *Origins of sound change: Approaches to phonologization,* 201–227. Oxford: Oxford University Press.