Phonetic evidence on phonology–morphosyntax interactions: Sibilant voicing in Quito Spanish

PATRYCJA STRYCHARCZUK, MARIJN VAN 'T VEER, MARTINE BRUIL and KATHRIN LINKE

Journal of Linguistics / *FirstView* Article / February 2014, pp 1 - 50
DOI: 10.1017/S0022226713000157, Published online: 10 July 2013

J. Linguistics, Page 1 of 50.
doi:10.1017/S0022226713000157

Phonetic evidence on phonology–morphosyntax interactions: Sibilant voicing in Quito Spanish[1]

PATRYCJA STRYCHARCZUK

University of Manchester

MARIJN VAN 'T VEER

Leiden University

MARTINE BRUIL

Leiden University

KATHRIN LINKE

Leiden University & The Meertens Institute, Amsterdam

(Received 7 August 2012; revised 8 April 2013)

This paper presents new experimental data on Quito Spanish /s/-voicing, which has attracted considerable interest from theoretical phonologists owing to the overapplication of voicing to word-final pre-vocalic /s/. Bermúdez-Otero (2011) singles out Quito /s/-voicing as an important test case for discriminating between two competing theories of phonology–morphosyntax interactions: Output–output correspondence and cyclicity. Overapplication in /s/-voicing cannot be captured using correspondence relationship to a base form, which challenges Output–output correspondence as a theory of opacity. However, the argument only holds insofar as word-final prevocalic /s/-voicing is considered phonological, as Output–output correspondence can account for /s/-voicing assuming that it only applies in the phonetics (Colina 2009). We discuss the diverging empirical predictions concerning categoricity and gradience in the surface realisation of voicing processes. We further test these predictions based on acoustic data from seven speakers of Quito Spanish. Evidence from speech rate manipulations shows that some speakers produce more voicing during frication at normal speech rate, compared to fast, maintaining a stable voicing ratio across different speech rates. We argue that for these speakers, /s/-voicing is optional but categorical, and so it ought to be analysed as phonological. This result presents a

[1] We would like to thank the speakers for their participation in the experiment. We are also grateful to Ricardo Bermúdez-Otero, Yuni Kim, Koen Sebregts and three anonymous Journal of Linguistics referees for their comments and suggestions. The research reported on in this article has been made possible thanks to a doctoral grant from the Arts and Humanities Research Council (AHRC, www.ahrc.ac.uk) to the first author, a grant from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), project number 360-75-000, to the second and forth author, and a grant from the NWO, project number 360-70-320, to the third author.

challenge to the Output–output correspondence approach, but can be accommodated within cyclicity.

1. Introduction

One of the basic challenges in research on phonological opacity arises from the lack of consensus concerning the boundary between phonology and other components of grammar. If theories of phonological opacity are limited to phonological misapplication, then their validity can only be tested based on cases which are themselves phonological in nature. The case of pre-vocalic /s/-voicing in Quito Spanish illustrates this issue, as competing explanations for its overapplication hinge on whether or not /s/-voicing is a phonological process at all, or whether it is more accurately described as phonetic. The current paper approaches the problem of the phonological or phonetic status of Quito /s/-voicing from an empirical perspective, and uses phonetic evidence to discriminate between a cyclic and correspondence-based approach to opacity.

The opacity problem in the case of Quito /s/-voicing concerns over-application of coda sibilant voicing before sonorants to word-final pre-vocalic sibilants which surface as onsets, for instance in cases like *gas acre* [ga.za.kɾe] 'acrid gas'. Bermúdez-Otero (2011) argues that this over-application cannot be explained as conditioned by correspondence to another surface form, as there is no appropriate base form with a word-final voiced variant. According to Colina (2009), however, the challenge is only apparent, assuming that a phonological output correspondence relationship can hold between two members of a paradigm that are phonetically quite distinct. Colina proposes that word-final pre-vocalic sibilants in Quito Spanish, as well as all pre-sonorant sibilants, undergo phonetic, but not phonological voicing. At the level of phonology, the argument goes, the /s/-voicing tokens are delaryngealised archiphonemes by correspondence to word-final pre-pausal sibilants. We scrutinise this claim in the current paper based on the assumption that any apparent phonetic process operating on top of phonological forms must show properties of phonetic gradience. This interpretation follows from a modular view of the phonetics–phonology interface, where gradient phonetic interpretation is applied to distinct phonological categories. Under this classical view, Colina's (2009) analysis predicts the existence of gradient effects in Quito Spanish /s/-voicing.

Using acoustic data from Quito Spanish speakers, we evaluate whether there is gradience in the /s/-voicing sequences. We argue that, at least for some speakers, /s/-voicing is realised as an optional but categorical process, rather than as gradient coarticulation. Evidence for this is the effect of speech rate on the acoustic exponents of voicing. The participants who show evidence for the existence of two categories in the /s/-voicing environment produce more phonetic voicing at normal, as opposed to fast, rate. Since less coarticulation is expected at slower speech rates, the increase in voicing duration suggests the existence of a voicing target. Based on this evidence we argue that a categorical interpretation involving optional feature assignment captures the existing data best.

There are two aims which we intend to achieve with this paper. First, we present original data on Quito pre-sonorant voicing, a process which has not as yet received a detailed acoustic description. Second, we discuss the consequences of our findings for earlier theoretical claims concerning the role of phonological opacity in Quito /s/-voicing.

1.1 *Quito /s/-voicing*

The variety of Spanish spoken in Highland Ecuador, including Quito, has been reported by Robinson (1979) and Lipski (1989) to voice word-final /s/ when followed by a vowel in the next word. Another process affecting word-final sibilants, as well as other types of final codas, is resyllabification into an onset. In the word-medial position pre-vocalic sibilants surface as voiceless. The relevant examples are in (1).

(1) *Pre-vocalic /s/ in Quito Spanish*[2]

/gas#akɾe/	[ga.za.kɾe]	‘acrid gas’
/gasita/	[ga.si.ta]	‘gauze (DIM)’

In addition, similarly to other varieties of Spanish, the Quito dialect has coda /s/-voicing before sonorants and before voiced obstruents, as shown in (2).

(2) */s/ before sonorants and voiced obstruents*

/plasma/	[plaz.ma]	‘plasma’
/gas#noble/	[gaz.no.βle]	‘noble gas’
/rasgo/	[raz.ɰo]	‘feature’
/gas#blaNko/	[gaz.βlaŋ.ko]	‘white gas’

In pre-pausal position and before a voiceless obstruent a coda /s/ surfaces as voiceless, as exemplified in (3).

(3) */s/ word-finally and before voiceless obstruents*

/gas/	[gas]	‘gas’
/gas#kaɾo/	[gas.ka.ɾo]	‘expensive gas’

1.2 *Opacity*

The unique and theoretically complex aspect of the Quito system is the positional restriction on pre-vocalic /s/-voicing; pre-vocalic voicing applies in

[2] All the data in this section come from Bermúdez-Otero (2011).

word-final position, but it fails to apply to word-medial tokens. Another way of phrasing this generalisation is in terms of syllable structure, as voicing affects derived word-initial onsets, but not word-medial onsets. Even though word-final resyllabification is not obligatory (it might be blocked in slow speech, or where there is glottal insertion, see Lipski 1989), it applies in most cases. This generalisation is supported by numerous descriptions (Harris 1969, 1983, inter alia), and by restrictions on some phonological processes. For example, emphatic trilling (Harris 1983) is blocked from applying to word-final rhotics followed by a vowel, e.g. *mar Egeo* [ma.ɾe.xe.o], but *[ma.re.xe.o] 'Aegean Sea', while it may optionally apply to rhotics in canonical coda positions, e.g. *mar* [maɾ] ~ [mar] 'sea'; *mar Negro* [maɾ.ne.ɰɾo] ~ [mar.ne.ɰɾo] 'Black Sea'.

Interestingly, the syllabification presented in (4) is disputed by Robinson (2012) for the case of Ecuadorian Spanish.

(4) *Contextual realisation of /n/ in Quito Spanish* (Harris 1983, Robinson 2012)

	Harris's syllabification	Robinson's syllabification	
/bien/	[bieŋ]	[bieŋ]	'good'
/bienes/	[bie.nes]	[bie.nes]	'goods'
/con/	[koŋ]	[koŋ]	'with'
/con#uno/	[ko.ŋu.no]	[koŋ.u.no]	'with one'

Robinson argues that while word-final resyllabification may be common in most varieties of Spanish, Ecuadorian speakers show non-standard syllabification judgements, where a word-final pre-vocalic /s/ is considered to be a coda. Native speakers of Ecuadorian Spanish consulted by Robinson (including speakers from Quito) syllabified forms such as *has ido* 'you have gone' as [az.i.ðo], judging a syllabification like [a.zi.ðo] as incorrect. Robinson reports non-standard syllabification judgements also for word-final pre-vocalic /n/. The realisation of /n/ in this position is associated with [n] ~ [ŋ] allophony, which closely resembles the [s] ~ [z] allophony in terms of its distribution. /n/ undergoes velarisation in word-final position, also when followed by a vowel in the next word. In word-medial onsets, however, only alveolar [n] is found. Robinson reports that word-final resyllabification may apply in Ecuadorian dialects for other consonants. For instance, for *del Oro* 'from el Oro' two syllabifications were considered possible by Robinson's informants: [del.o.ro] and [de.lo.ro].

The issue of syllabification is crucial to the discussion of opacity in Quito /s/-voicing, as the generalisation that /s/-voicing overapplies in derived word-initial onsets is tied to the assumption that resyllabification does indeed apply. Assuming the syllabification judgements reported by Robinson, the

Stage	Change	Examples
1	Initial state	[as.i.ðo] [kon.u.no] [del.o.ro]
2	/s/-voicing and /n/-velarisation	[az.i.ðo] [koŋ.u.no] [del.o.ro]
3	Resyllabification	[az.i.ðo] [koŋ.u.no] [de.lo.ro]

Table 1
Hypothetical development of /s/-voicing as a sequence of transparent synchronic systems.

application of Quito /s/-voicing could be formulated as the following transparent generalisation: the voicing applies to pre-vocalic and pre-sonorant coda /s/. However, contrary to what Robinson (2012) concludes, the reported native syllabification judgements do not establish that the opacity problem is merely apparent. First, it is not clear what exactly conditions the judgements of native speakers in a metalinguistic syllabification task, such as the one employed by Robinson. If native speakers judge a syllabification such as [a.zi.ðo] for *has ido* 'you have gone' as unacceptable, it may not necessarily be due to the actual syllable structure, as other, for instance lexical, factors may also be involved. Voiced fricatives are never found in canonical onsets in Quito Spanish, which may influence native speakers' perception of how well-formed a syllable like [zi] is. This objection is particularly valid considering that the only other case where resyllabification is reported to be blocked is the case of word-final pre-vocalic /n/, where just as in the case of /s/, a morphosyntactically conditioned segmental process creates pre-vocalic velar nasals, which are not found in the canonical onset position.

The second argument for why the opacity problem in Quito Spanish requires a systematic analysis comes from diachrony. Even if one assumes that word-final /s/ and /n/ resist resyllabification in Quito Spanish, the question arises whether the diachronic development of these two segmental processes could have involved a series of transparent synchronic grammars. This could only be possible if resyllabification were shown to diachronically follow /s/-voicing and /n/-velarisation, as schematised in Table 1. Under this scenario, word-final pre-vocalic consonants all surface in codas at stage 1. At stage 2, /s/-voicing and /n/-velarisation apply transparently in codas, including word-final pre-vocalic codas. At stage 3, resyllabification applies to word-final pre-vocalic consonants, with the exception of /s/ and /n/.

However, while the transparent diachronic scenario sketched out above is in principle possible, there is no independent evidence to support it. It is difficult to precisely date the relative development of /s/-voicing, /n/-velarisation and resyllabification in Quito Spanish based on the existing sources, but resyllabification is likely the relatively oldest process. One source of evidence for this claim is the fact that resyllabification applies in a wide range of both Peninsular and South American Spanish dialects, whereas /n/-velarisation and /s/-voicing are geographically more restricted. Furthermore, there is evidence that resyllabification was already operative in Latin. Ryan (forthcoming) discusses the case of prosodic word minimality in Latin, and its interaction with syllabification. Prosodic words in Latin are minimally bimoraic (CVː or CVC). Thus, pre-vocalic CVC words are potentially made subminimal by post-lexical resyllabification, losing their final mora. In an analysis of Virgil's *Aeneid* and *Georgics*, Ryan finds avoidance of monosyllabic content CVC words in the pre-vocalic position. No such tendency is observed for monosyllabic CVːC words, where minimality requirements are not violated by resyllabification. Similarly, word-final CVC sequences are widely attested in the corpus for polysyllabic words followed by a vowel. On the basis of these findings, Ryan argues that Virgil avoids monomoraic feet created by post-lexical resyllabification.

If resyllabification did indeed predate /s/-voicing, then /s/-voicing most likely developed as an opaque process.[3] Even if resyllabification was subsequently blocked for word-final /s/ and /n/, making the two processes transparent, an earlier opaque stage also needs to be accounted for. In our analysis, we assume resyllabification for word-final /s/, as we consider resyllabification to be a possible rule of present day Quito Spanish (contra Robinson 2012), or at the very least a rule of a diachronically earlier stage of this dialect.

Resyllabification of word-final voiced pre-vocalic sibilants into onsets creates a complex phonological problem, especially to monostratal models of phonology such as Classic Optimality Theory (OT; Prince & Smolensky 2004 [1993]). The crucial assumption of Classic OT is that phonological processes reflect interactions of constraints on phonological output. One consequence of this approach is that phonological generalisations must be surface-true. This is not the case with Quito /s/-voicing. The generalisation that coda /s/ undergoes voicing before a sonorant overapplies in the case of a final /s/ which surfaces in an onset.[4] One possible way of modelling Quito /s/-voicing in Classic OT is by relegating it from the domain of phonology, as proposed by Colina (2009). Colina argues that word-final sibilants undergo delaryngealisation, to satisfy the restriction against laryngeal licensing in the coda position. As a result, a coda /s/ surfaces as a delaryngealised archiphoneme S in word-final position, as illustrated in the tableau in (5).

[3] We return to the discussion of how a process like this might develop through a series of well-attested sound changes in Section 4.2 below.

[4] For a detailed discussion of how a Classic OT analysis fails to derive the Quito Spanish facts the reader is referred to Colina (2009).

(5) *Delaryngealisation of coda /s/ in order to satisfy the coda licensing restrictions*[5]
/los/ [los]

/los/			LICENSE[lar]	*S	*z	*s	IDENT(voice)
a.		loz	*!		*		*
b.		los	*!			*	
c.	☞	loS		*			*

The delaryngealised archiphoneme has no voicing target in the output of phonology, and thus its voicing is supplied by the phonetics, in accordance with Phonetic Underspecification Theory (Keating 1988, 1990). The voicing targets are supplied depending on the neighbouring segments. In final position, the default realisation will be voiceless, whereas when a laryngeally unspecified sibilant is surrounded by phonetically voiced segments, the voicing might spill over, yielding a voiced sibilant realisation.

The analysis so far accounts for the voiceless realisation of pre-pausal sibilants and sibilants followed by voiceless stops, and for voiced realisation of coda /s/ followed by sonorants and voiced obstruents. What remains to be explained is why the voicing applies in word-final sibilants followed by a word-initial vowel. Delaryngealisation is superfluous in these cases, as the final sibilant surfaces in an onset, vacuously satisfying the restriction against laryngeal licensing in codas. However, Colina (2009) proposes that in these cases the sibilant undergoes delaryngealisation as well, in order to satisfy a high-ranked Output–output constraint which requires identity to the base form (e.g. identity of *los* in *los otros* 'the others' to the base form *los*). The tableau in (6) provides an illustration.

(6) *Output–output conditioned delaryngealisation of a word-final /s/ followed by a vowel*
/los otros/ [lo.zo.tros] and [los]

/los otros/			IDENT-OO (voice)	LICENSE[lar]	*S	*z	*s	IDENT(voice)
a.		lo.zo.troS	*!		*	*		*
b.		lo.so.troS	*!		*		*	
c.	☞	lo.So.troS			**			*

[5] The constraints used in this tableau and the tableau in (6) below reflect Colina's (2009) analysis and are used for expository purposes.

A pre-vocalic delaryngealised S is expected to be realised with some voicing through interpolation from the neighbouring vowels, leading to voiced perceptions. Colina (2009) argues that the two mechanisms assumed in her analysis – phonetic underspecification and Output–output correspondence – explain the voicing patterns found in Quito Spanish, and account for the 'gradient and variable' voicing as observed in the language. However, as pointed out by Bermúdez-Otero (2011), Colina's (2009) analysis not so much explains the gradient and variable character of /s/-voicing, as crucially predicts it by virtue of assumed representations. The Output–output correspondence analysis presented above relegates /s/-voicing from the phonological domain due to theoretical rather than empirical considerations, as phonologically opaque interactions cannot be easily handled within a strictly monostratal approach.

Bermúdez-Otero (2011) argues that whether or not Quito /s/-voicing is a phonetic process is down to empirical evidence. Should empirical evidence show that /s/-voicing is indeed categorical, an Output–output relationship between phonetically underspecified members of a phonological paradigm would not suffice to fully account for the Quito Spanish facts. As a potential alternative which could handle categoricity in /s/-voicing, Bermúdez-Otero (2011) considers a derivational analysis based on lexical strata (Kiparsky 1982, 1985). Within this approach, the rule of /s/-voicing applies post-lexically. Bermúdez-Otero (2011) proposes that coda /s/ undergoes delaryngealisation at the Word Level (WL), conditioned by the inability of codas to license voicing. This process turns the sibilant into a delaryngealised archiphoneme S. This output is further submitted to the Phrase Level (PL) of derivation. At this level, the delaryngealised S is voiced before a vowel (or any other voiced sound), but interpreted as voiceless default before a pause. Resyllabification also operates at the Phrase Level. Crucial to the analysis is the fact that a word-medial pre-vocalic /s/, as in *gasa* 'gauze', is not a coda at any level of derivation. Consequently, it does not undergo coda delaryngealisation, and thus it does not meet the conditions for Phrase Level voicing, which only applies to the laryngeally unspecified S. The two types of derivation (for word-medial and word-final) /s/ are illustrated in (7).

(7) *A cyclic derivation of /s/-voicing in Quito Spanish* (Bermúdez-Otero 2011)

	$[_{\text{PL}}[_{\text{WL}}\text{gasa}]]$	$[_{\text{PL}}[_{\text{WL}}\text{gas}][_{\text{WL}}\text{akre}]]$	$[_{\text{PL}}[_{\text{WL}}\text{gas}]]$
Word Level (coda delaryngealisation)	[ga.sa]	[gaS][a.kɾe]	[gaS]
Phrase Level (assimilation and default)	[ga.sa]	[ga.za.kɾe]	[gas]

A rather different analysis of the Quito /s/-voicing pattern, couched within the Dispersion Theory (Flemming 1995, 2002), is provided by Bradley (2005) and Bradley & Delforge (2006). Within this analysis, /s/-voicing is understood as a functional process aimed at maintaining systemic contrasts. Bradley (2005) and Bradley & Delforge (2006) observe that Quito Spanish /s/-voicing is associated with contrast at the post-lexical level, as illustrated by the pair in (8).

(8) *Contrast created by /s/-voicing*
has ido [haziðo] 'you have gone'
ha sido [hasiðo] 's/he, it has been'

Surface contrasts are evaluated by systematic constraints which maximise the perceptual distinctness of output forms. The voicing of word-final pre-vocalic /s/ is formalised within this approach as a two-step process. Lexical phonology generates laryngeally underspecified sibilants in the word-final position and [–voice] sibilants in the onset position. The assignment of [–voice] in onsets is conditioned by a high-ranked constraint $_{\sigma}$[s, which requires onsets /s/ to surface as voiceless. In the coda position (at the lexical level), /s/ is mapped onto a delaryngealised archiphoneme S, due to markedness restrictions against [±voice] assignment. Surface /s/-voicing as found in Quito Spanish is attributed to two distinct mechanisms. The first of these mechanisms is responsible for pre-vocalic voicing, while the other results in pre-consonantal voicing. Pre-vocalic voicing is seen as conditioned by the constraint Space$_{SV}$, formulated in (9), which favours the maximal possible contrast between word-final and word-medial pre-vocalic /s/, in order to distinguish phrasal minimal pairs which occur in the language.

(9) Space$_{SV}$ (Bradley 2005, Bradley & Delforge 2006)
Potential minimal pairs differing in sibilant voicing differ at least as much as [s] and [z] do between vowels.

Word-final pre-vocalic /s/-voicing is seen as applying in post-lexical phonology, and is analysed as a categorical process. Pre-consonantal /s/-voicing, on the other hand, is analysed as phonetic. Post-lexical phonology generates the delaryngealised archiphoneme S in all codas (which is, again, due to markedness restrictions against laryngeal specifications in obstruents). This underspecified segment is then subject to coarticulatory gradient voicing in the phonetics.

The diverging assumptions concerning the mechanisms that condition /s/-voicing made by the analyses presented above lead to diverging empirical predictions in the case of Quito Spanish /s/-voicing. Colina's (2009) analysis crucially predicts all cases of /s/-voicing to be gradient, whereas the cyclical account put forward by Bermúdez-Otero (2011) can accommodate categorical /s/-voicing. The analysis by Bradley (2005) and Bradley & Delforge (2006) makes the prediction that /s/-voicing is gradient pre-consonantally, but categorical in the word-final position before a vowel. Therefore, the predictions of the different theories can be tested by means of instrumental phonetic data.

1.3 *Gradience and categoricity*

The predictions of the analyses presented above follow straightforwardly from the modular view of the phonetics–phonology interface, which all the authors assume (if an analys allows an exclusion of /s/-voicing from the phonological domain, it follows that phonology is considered distinct from phonetics). Under the modular view of the interface, along the lines of Keating (1996), phonology is categorical and abstract, whereas phonetics realises abstract phonological categories in the continuous physical dimension. Speech consists of both phonological and phonetic phenomena, crucially distinguished by gradience. While phonological operations affect category labels, phonetic effects are gradient, i.e. they consist in 'gradual changes over time along quality dimensions' (Keating 1996: 263). Whether a given process is phonetic or phonological is thus an empirical, not a theoretical question, as argued by Myers (2000) and Cohn (2006), among others. It is also a question that has guided a whole line of experimental research in laboratory phonology, including Barry (1992), Cohn (1993), Holst & Nolan (1995), Zsiga (1995), Nolan, Holst & Kühnert (1996), Ellis & Hardcastle (2002) and Tucker & Warner (2010).

The existing data on Quito /s/-voicing, based on discrete coding, do not allow us to conclude whether the process is categorical or gradient. Robinson's (1979) data come from interviews which he conducted during fieldwork in Ecuador. The participants read a word list containing tokens of pre-vocalic /s/ in different positions with respect to word and morpheme boundary. The tokens were then auditorily classified by the investigator as either voiced or voiceless, and transcribed. Robinson's description was later supplemented by Lipski's fieldwork which was also based on auditory transcriptions (see Lipski 1989). Pre-vocalic /s/-voicing in Quito Spanish has also been recently studied by Chappell (2011), who recorded 404 tokens of pre-vocalic /s/-voicing in different prosodic and morphological environments. Chappell's (2011) data come from Quito radio stations' archives, and feature local speakers discussing local issues and events. The data were analysed acoustically and labelled as voiced based on the presence of a strong voicing bar. If no such bar was present, the tokens were categorically labelled as voiceless.

Despite the lack of conclusive instrumental evidence, some researchers working on Quito /s/-voicing have considered the process gradient. Lipski (1989: 55) comments that the /s/-voicing is 'variable and gradient'. Chappell (2011: 63) notes that /s/-voicing at word boundaries in Quito Spanish is 'not as categorical as previously claimed'. However, although neither Lipski (1989) nor Chappell (2011) are explicit about this, they seem to use 'gradience' as a cover term for both quantitative and qualitative variation. This kind of ambiguous terminological approach to gradience is sometimes used in the literature, but the difference is crucial. Chappell (2011) presents firm evidence of variability of /s/-voicing, but not of phonetic gradience. In her study, the presence of a voicing bar was recorded for 98% of word-final sibilants followed by a vowel in the next word. When a sibilant was followed by a vowel in the same word, the voicing bar was typically absent; this was the case for 89% of /s/+vowel sequences word-medially, and in 94% of such sequences at the beginning of a word. Since voicing was variably realised in different environments, the numbers confirm that there was optionality. What they do not tell us is whether there were cases of partial voicing, where the voicing bar was present during just some part of the sibilant, which would be symptomatic of a phonetically gradient process. Due to these issues, we distinguish between gradience and optionality in our discussion, following Pierrehumbert (2006), who also points out that empirical methods used for identifying phonetic gradience critically rely on the use of continuous phonetic data.

1.4 *The current study*

Different theoretical approaches to phonological opacity give rise to crucially diverging empirical predictions regarding the categorical or gradient status of /s/-voicing in Ecuadorian Spanish. Since these predictions are impossible to evaluate on the basis of the discrete data found in the literature to date, we have collected new continuous phonetic data, which we analyse with a view to establishing whether Quito /s/-voicing is categorical or gradient.

We base our diagnostics of categoricity on a number of methods emerging from the laboratory phonology literature. We begin our investigation with a qualitative spectrographic analysis of the acoustic realisations of /s/ in a number of phonetic contexts. Voicing in fricatives is not lexically contrastive, but voiced and partially voiced fricative allophones are found in all varieties of Spanish, and they are distinguished by the presence of vocal fold vibration (Williams 1977, Hualde 2005). Thus, effects of categoricity or gradience in Quito Spanish /s/-voicing are expected to be observed in the degree of vocal fold vibration during fricative production. Jansen (2004) and Ernestus & Baayen (2006) report varying durations of vocal fold vibration in clusters of obstruents followed by a voiced stop, which Ernestus (2011) cites as evidence for a gradient rule of voice assimilation in Dutch. We follow Jansen and Ernestus in a qualitative and quantitative evaluation of the duration of vocal fold vibration as an indicator of categoricity and gradience. However, we are careful to evaluate the evidence for gradience in a population against gradience at the level of individual speakers. The currently available empirical evidence on coarticulation and assimilation phenomena suggests that they are commonly characterised by variability, but that the variability can be either of gradient type or optional but categorical type. Data from Ellis & Hardcastle (2002) on /n#k/ sandhi in English provide an illustration of different types of variation in place assimilation. Two out of the ten participants in the Ellis & Hardcastle study produced a continuum of realisations from a fully executed coronal closure to no tongue-tip raising for the underlying /n/ in sequences like *ban cuts*. Two other speakers, on the other hand, varied their pronunciation between a fully realised coronal gesture for /n/ (no coarticulation at all), and full dorsal realisation with no residual coronal gesture (complete assimilation). These two different types of variation indicate that gradience and optionality need to be distinguished at the level of description, and that intra-speaker variation alone does not entail gradience. In order to evaluate the status of variation, one needs to address the issue of how the variable realisations are distributed for individual speakers.

The key test we employ for distinguishing gradient and categorical phenomena is based on speech rate manipulation, following Solé (1992, 1995). Solé uses manipulations of speech rate as diagnostics for distinguishing between mechanical phonetic properties which follow automatically from the vocal tract constraints and controlled speech processes. While mechanical properties are not expected to be affected by variations in speech rate, the acoustic cues that are under speaker's control undergo durational adjustments to maintain a steady ratio to the segment duration across different speech rates. Vowel nasalisation in Spanish and American English provide an illustration. Solé (1992, 1995) analysed the correlation of speaking rate and the duration of the nasal portion of a vowel followed by a nasal consonant. In Spanish, the nasalised vowel portion retained a relatively stable duration, and did not correlate significantly with the speaking rate. In American English, on the other hand, the duration of nasalisation varied systematically with the speaking rate, while the ratio of nasal portion to the total vowel duration remained stable at 100%. Solé proposes that the short period of nasalisation observed in Spanish is automated and it follows from the velum lowering produced in the anticipation of the nasal consonant. However, this kind of overlap is insufficient to explain the extent of rate-independent nasalisation in American English. Thus, for American English, Solé argues that vowel nasalisation is speaker-controlled and actively targeted. The speech rate test is also applied by Cuartero Torres (2001) to voice assimilation in Catalan and English.

In our study, we use the speech rate test to compare the effect of speech rate manipulation on different phonetic exponents of voicing, including the duration of vocal fold vibration, as well as the effect of speech rate on the voicing ratio. If Quito /s/-voicing is a gradient phenomenon, and the observed voicing is due to gestural overlap with the neighbouring voiced sound, we expect to find approximately the same voicing duration across different speech rates. Additionally, we expect to observe considerable differences in the ratio of voicing to fricative duration depending on speech rate, as the ratio will increase with a decrease in segment duration brought about by faster speech. If /s/-voicing is categorical, however, the voicing ratio is expected to remain relatively stable across different speech rates, in which case the voicing duration varies proportionally to the variation in segment duration.

We also consider how voicing duration and voicing ratio are affected by speaker sex. In a review of a number of studies of voicing, Jessen (2009) finds that male speakers typically show greater voicing duration and ratio than female speakers. According to Jessen, this difference can be explained in terms of physiological differences. The production of voicing conflicts with the rise in supraglottal air pressure, associated with the production of oral stricture. Therefore, vocal fold vibration is naturally impeded in obstruents (Ohala 1983, Westbury & Keating 1986). Jessen observes that intraoral pressure is expected to rise faster in smaller oral cavities, leading to quicker cessation of voicing. All things being equal, females are expected to show relatively shorter voicing duration than males. Importantly, speakers may execute an array of voicing gestures to prolong vocal fold vibration throughout stricture, overriding the physiological effects to an extent. Examples of voicing gestures include lowering the larynx, raising the soft palate and advancing the tongue (Ladefoged 1973, Stevens 1998). Jessen (2009) argues that effects of speaker sex on voicing duration are most readily observable in cases where no additional voicing gestures are expected, i.e. when the voicing is passive. Building on this argument, Jessen interprets speaker sex effects on intervocalic stop voicing in German as evidence for the voicing being passive. Extending Jessen's analysis to Quito Spanish, a significant effect of speaker sex on the duration of voicing and voicing ratio in Quito Spanish /s/-voicing could support a gradient interpretation of the process. Analysing the effect of speaker sex on voicing is further motivated by the findings on /s/-voicing in other dialects of Spanish. In a study of intervocalic /s/-voicing in Loja, Ecuador, García (2011) found that male speakers were more likely to produce fully voiced sibilants than females. Schmidt & Willis (2011) report significant interactions between the speaker sex and the prosodic and segmental context in the voicing of /s/ in Mexico City. We test the predictions concerning speech rate and speaker sex on the Quito Spanish data, paying close attention to how the result is related to individual variation.

2. Materials and method

2.1 *Stimuli*

The test items included sequences of word-final /s/ followed by a vowel in the next word, as well as of /s/ followed by a sonorant consonant within the same word and across a word boundary (we chose nasals in all cases). In the remainder of the paper we will be referring to these environments as the Quito /s/-voicing contexts. Even though it is only the word-final

	Word-internal		Word-final	
Pre-vocalic	1. *gasita*	'gauze (DIM)'	3. *gas acre*	'acrid gas'
Pre-sonorant	2. *entusiasmo*	'enthusiasm'	4. *gas noble*	'noble gas'
Pre-voiced obstruent	5. *esbozo*	'plan, sketch'	7. *marchas buenas*	'good marches'
Pre-voiceless obstruent	6. *obispo*	'bishop'		—

Table 2
Example test and baseline items. The baseline items are indicated by shading.

pre-vocalic /s/-voicing that distinguishes Quito Spanish from other Spanish dialects, the pre-vocalic and the pre-sonorant environments are similar in that they involve voicing before a segment which is not itself contrastively voiced. Examples of the test items are in Table 2, while the full list of test items is in Appendix A. There were six tokens per condition.

In addition to the test items, we studied the realisation of /s/ in a range of baseline contexts. The Quito dialect, like most other varieties of Spanish, does not have a lexical voicing contrast in sibilants (although contrast is present at the postlexical level in Quito Spanish, as shown in (8) above). In order to establish a baseline we decided to compare the realisation of /s/ in the test items to the realisation of /s/ before a voiced and voiceless stop within the same word. The latter condition involved only five items (as opposed to six) due to difficulties in finding items of comparable word size. In addition, we included 6 items of word-final /s/ followed by a voiced obstruent in the following word, as well as six items of a word-medial pre-vocalic /s/.

All the test and baseline items were embedded within a carrier phrase, as exemplified in (10).

(10) *The carrier phrase*
Diga **'gas acre'** otra vez.
'Say "acrid gas" one more time.'

2.2 *Participants*

Eight speakers participated in the experiment: four males aged between 16 and 25 years, and four females aged between 21 and 28 years. They were all born in Quito, where they also lived at the time the experiment was conducted. None of the speakers reported a history of speech or hearing deficiency. The data from one of the speakers (speaker 6) were discarded, as she produced multiple pauses and hesitations, and was judged to be markedly less fluent than others during the analysis of the recordings.

2.3 *Procedure*

The recordings were made in a quiet room on a Marantz PMD 620 solid state recorder, with a Sennheiser PC 131 headset. The sampling rate was 48 kHz.

Four repetitions were recorded per speaker, and each repetition was followed by a short break. The stimuli were randomised for each repetition, and presented to the speaker one at a time on a computer screen. For the first two repetitions the speakers were instructed to read at a comfortable pace, and encouraged to correct themselves if they were unhappy with their pronunciation. For the second two repetitions the participants were asked to say the sentences fast, as if they were saying them to someone who was about to leave the room. Altogether 41 (test items) × 4 (repetitions) × 7 (speakers) = 1148 tokens were collected and analysed acoustically. One hundred and ten tokens (9.5% of the data) were discarded due to disfluencies, reading errors, glottalisations, etc., leaving 1048 tokens for statistical analysis.

2.4 *Acoustic analysis*

For the analysis of voicing we measured the duration of vocal fold vibration during frication (henceforth: voicing duration), fricative duration and intensity difference between low- and total-frequency intensity (henceforth: intensity difference).

An extended period of vocal fold vibration is a common voicing cue in fricatives. The degree of fricative voicing can be quantified in terms of the duration of vocal fold vibration. This approach has been previously used by Jansen (2004) and Ernestus & Baayen (2006) to quantify the degree of gradient voicing in Dutch (recall Section 1.4 above), as well as by Chappell (2011), who reports that varying degrees of voicing duration are found in the /s/-voicing context in Ecuadorian Spanish.

The voicing contrast in fricatives is cued in some languages by the duration of the fricative. Voiced fricatives have been found to surface as shorter than voiceless ones in a number of languages, including Dutch (Slis & Cohen 1969a) and English (Crystal & House 1988). We measured the duration of fricatives in various voicing contexts in our data to verify whether there is an effect of voicing on the fricative duration in Ecuadorian Spanish. Measurements of fricative duration were also used to calculate the ratio of voicing to frication, with a view to assessing the effect of speech rate on voicing duration as opposed to voicing ratio.

As an additional potential exponent of voicing, we considered the difference between low- and total-frequency intensity. Gradoville (2011) reports that gradient differences in voicing in Argentinian Spanish are reflected in fricative intensity. We followed Gradoville's method in calculating the intensity difference which also serves as a normalisation strategy. We measured the mean intensity of the low-frequency portion of the fricative (filtered from

0 Hz to 900 Hz) and the mean intensity of the unfiltered fricative. We then subtracted the total intensity from the low-frequency intensity to calculate the intensity difference. As all the resulting values were negative, we refer to the absolute value of the intensity difference in the rest of this paper.

We also measured the duration of the latter portion of the carrier sentence (the duration of the phrase *otra vez* 'one more time') for each sibilant token in order to compare the speaking rate across different speakers, as well as quantify the extent of experimental speech rate manipulation on the speakers' pronunciation.

All the measurements were based on manual segmentation performed in Praat (Boersma & Weenink 2009) by one of the authors. Fricatives were identified based on the presence of high-frequency noise. The duration of voicing during frication was identified based on the presence of the voicing bar on the spectrogram and the presence of periodicity in the waveform. An accuracy check of the annotation was performed on 100 randomly selected tokens by another experimenter. For 95% of tokens, there was an absolute agreement concerning the annotation. For the remaining 5% of tokens, there was some discussion concerning the offset of partial voicing, where the voicing was fading out gradually during frication. It was agreed that in such cases the offset of voicing would be placed where the final pulse was detected by the Praat pitch tracker. The data in the entire corpus were re-analysed accordingly. The Pitch range was set to 75–300 Hz for male speakers, and to 100–500 Hz for female speakers. Intensity was measured with a Praat script based on the manual annotation.

3. Results

3.1 *Phonetic realisation*

Initial exploration of the phonetic data reveals that there is a considerable degree of inter- and intra-speaker variation with respect to the duration of vocal fold vibration during frication in the /s/-voicing context. Individual pronunciations of /s/-voicing fricatives ranged from voicing present throughout the fricative to realisations with very limited voicing. The left panel in Figure 1 provides an example of a fully voiced realisation with voicing present throughout frication. This kind of complete voicing was found in 60% of tokens within the /s/-voicing context. The remaining cases of /s/-voicing were typically realised with partial voicing. For the /s/-voicing cases where the voicing did not extend throughout frication, the median voicing duration was 18.59 ms and the median voicing ratio was .29.[6] The right panel of Figure 1 illustrates a partially voiced realisation.

[6] The classification of voicing as full or partial is used here to give the reader an overview of the variation we find, but it does not represent a systematic analytical approach to the data.

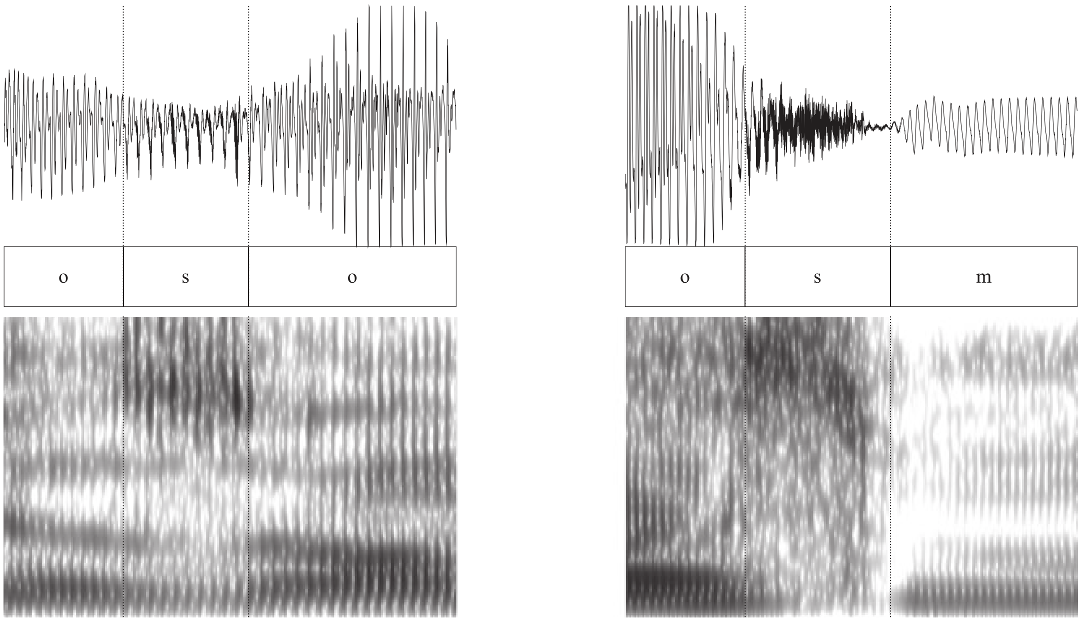

Figure 1
Left: Fully voiced /s/ in *muchos hombres* 'many men' pronounced by Speaker 1. Right: Partially voiced /s/ in *cosmologo* pronounced by Speaker 5.

Similarly to the /s/-voicing contexts, there was some variation in the realisation of voicing in coda sibilants followed by a voiced stop (either in the same word, or across a word boundary). Full voicing extending throughout frication was found in 72% is such cases. An example of a fully voiced sibilant followed by a voiced stop is in the left panel of Figure 2. The remaining cases of coda sibilants followed by voiced stops were typically realised with partial voicing, with median voicing duration of 18.51 ms and median voicing ratio of .28. The right panel of Figure 2 shows an example realisation.

Coda sibilants followed by a voiceless stop and sibilants in word-medial onsets were typically realised as either fully voiceless, or as partially voiced. Completely voiceless realisations were found in 34% of tokens in this context. An example realisation is illustrated in the left panel of Figure 3. When partial voicing was found, it was typically of limited duration, with the median voicing duration of 11.6 ms and median voicing ratio of .14. An example of a partially voiced sibilant in a word-medial onset is in the right panel of Figure 3.

Consequently, the figures quoted in this section should not be treated as conclusive statistics about the extent of Quito /s/-voicing.

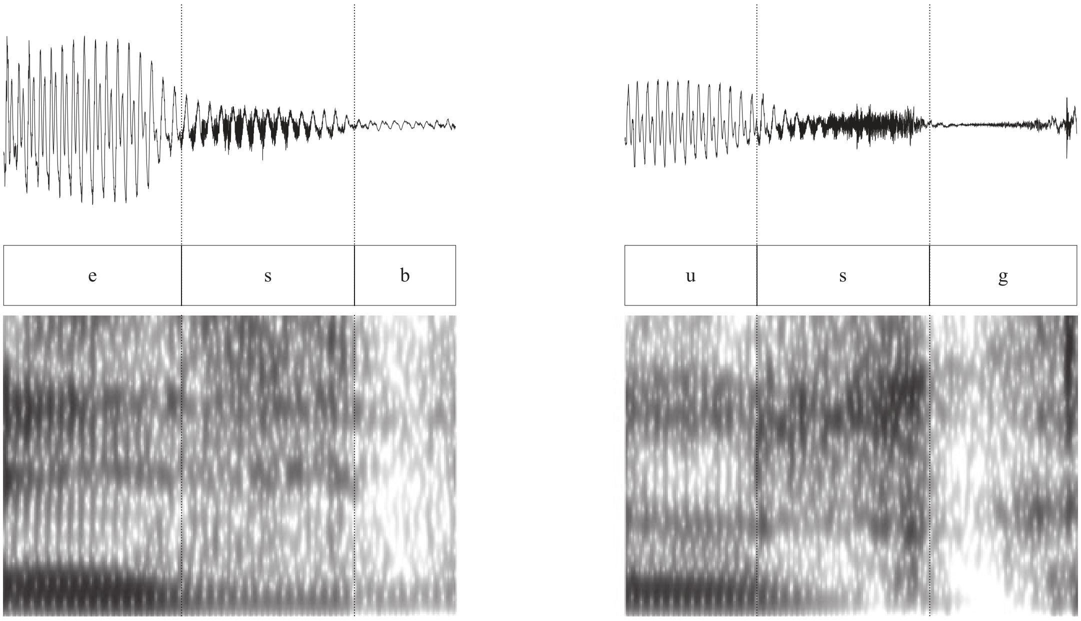

Figure 2
Left: Fully voiced /s/ in *presbítero* 'priest' pronounced by Speaker 5. Right: Partially voiced /s/ in *cactus grande* 'big cactus' pronounced by Speaker 5.

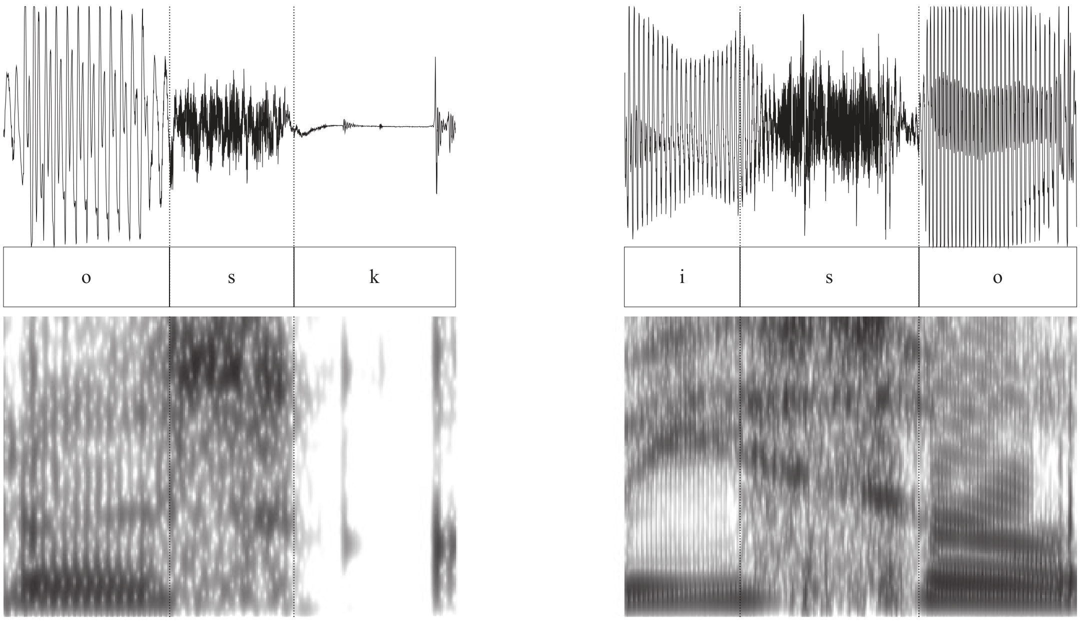

Figure 3
Left: No voicing of /s/ in *microscopio* 'miscroscope' pronounced by Speaker 5. Right: Partially voiced /s/ in *aviso* 'notice' pronounced by Speaker 5.

3.2 *Statistical analysis*

3.2.1 *Quito /s/-voicing contexts vs. other environments*

The first step in the statistical analysis involved exploring how various acoustic correlates of voicing are realised depending on the prosodic and the segmental environment. Given the scarcity of experimental evidence on Quito Spanish /s/-voicing, our aim was to verify the empirical status of /s/-voicing before sonorants and pre-vocalically in the external sandhi context, compared to sibilant voicing elsewhere. We analysed four acoustic correlates of voicing: voicing duration, voicing ratio, duration of the sibilant, and the intensity difference, using conditional inference trees. The analysis was performed in R (R Development Core Team 2005), version 2.13.1, using the party package (Hothorn, Hornik & Zeileis 2006). The raw means and standard deviations for the four dependent variables depending on three predictors are provided in Appendix B.

Conditional inference trees are a non-parametric recursive partitioning technique. The algorithm selects the predictor that provides the best binary split for the data, performs the split, and repeats the procedure until no further significant splits can be made. The output is a regression tree which provides a visualisation of how the value of the dependent variable can be predicted. The tree branches according to significant splits, and it also includes boxplots of distributions of the dependent variable for the terminal nodes. The recursive aspect allows for insight into the hierarchical structure of predictors, whereas the binary splits allow one to identify clusters of data. This set of properties makes conditional inference trees a useful tool for discovering structure within the data. Although this statistical method does not appear to have been widely used in phonetic research (as far as we have been able to ascertain), it has been shown to provide useful insights in modelling language variation (Tagliamonte & Baayen 2012). Another important advantage of this technique is that it makes no assumptions about the distribution of the response variables, and so it is applicable to analysing non-normally distributed data.

3.2.1.1 *Voicing duration*

Figure 4 presents a conditional inference tree for voicing duration measured in ms based on a model with three predictors: condition, rate and speaker sex. The first split separates Conditions 1 and 6 from the remaining contexts. Condition 1 involved word-medial onsets (*casita*), while Condition 6 involved sibilants followed by voiceless stops (*obispo*). Thus, the split confirms that these two conditions are characterised by significantly shorter vocal fold vibration compared to sibilants in the remaining contexts. The remaining contexts span the /s/-voicing cases, as well as sibilants followed by voiced stops. Within this group there is a significant split conditioned by speech rate, with relatively shorter voicing duration within the fast speech rate condition. Within the normal speech rate conditions there is a further split separating Conditions 5 and 7 from Conditions 2, 3 and 4. Conditions 5 (*esbozo*) and 7 (*marchas buenas*), where the voicing duration was relatively greater, had a voiced stop following a sibilant. Conditions 2, 3 and 4 are the

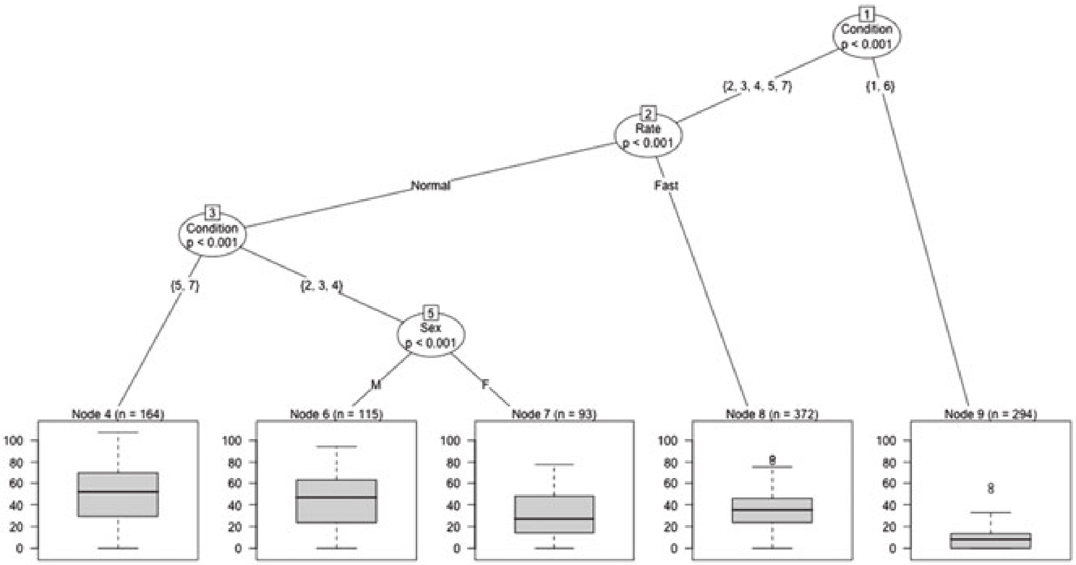

Figure 4
Conditional inference tree for voicing duration (in ms) based on a model with three predictors: condition, speech rate and speaker sex.

/s/-voicing conditions, involving sibilants followed by sonorants (Condition 2: *entusiasmo*, Condition 4: *gas noble*) or a vowel in the following word (Condition 3: *gas acre*). The model confirms that the /s/-voicing conditions cluster together with respect to voicing duration, since no further significant splits were made separating Conditions 2, 3 and 4. However, there was a further split within the /s/-voicing cases conditioned by speaker sex, with male speakers showing significantly greater voicing duration than female speakers.

3.2.1.2 *Voicing ratio*

Figure 5 presents a conditional inference tree for the ratio of voicing duration to the duration of the sibilant. The results are very similar to those of the model of voicing duration as far as /s/-voicing is concerned. The first split separates the /s/-voicing contexts as well as sibilants followed by voiced stops from those sibilants where no voicing is expected, i.e. word-medially before a vowel (Condition 1: *casita*) and before a voiceless stop (Condition 6: *obispo*). The further splits in the left branch of the tree are identical to those in the case of voicing duration. The fast speech rate condition was characterised by an overall higher ratio, and there were no further splits depending on condition within the speech rate. Within the normal rate, however, there was a split between sibilants followed by voiced stops (Condition 5: *esbozo*, Condition 7: *marchas buenas*) and the /s/-voicing environments (Conditions 2, 3 and 4). The former were characterised by a relatively higher voicing ratio compared to the latter. The /s/-voicing conditions were again found to cluster together, with no further splits. An effect of speaker sex was observed within

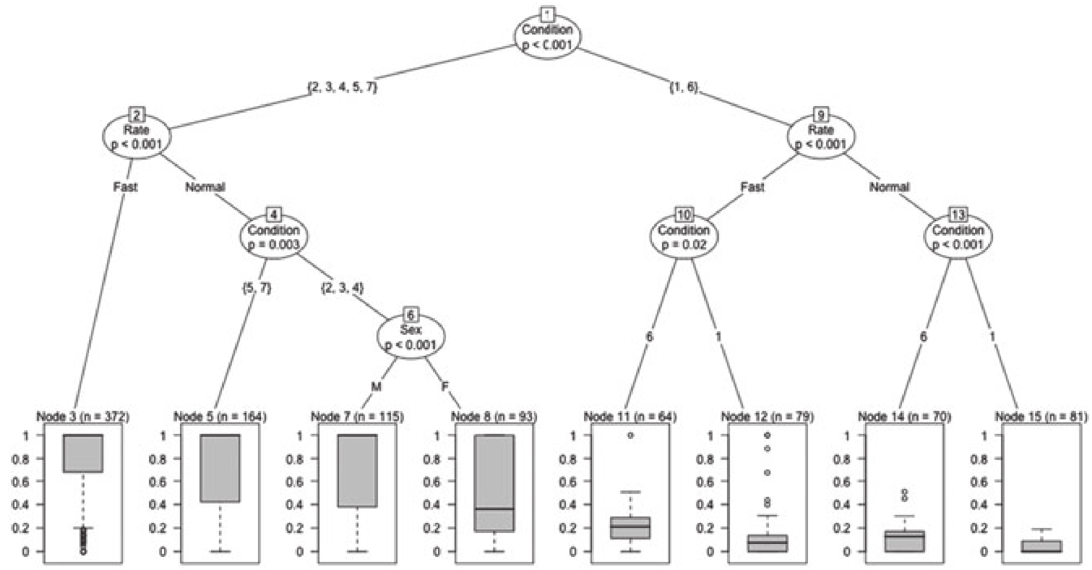

Figure 5
Conditional inference tree for voicing ratio based on a model with three predictors: condition, speech rate and speaker sex.

the /s/-voicing conditions only, with a higher voicing ratio in male speakers compared to female speakers.

3.2.1.3 *Fricative duration*

A tree for fricative duration (Figure 6) represents a different aspect of the data structure. The first split singles out Condition 1 (*casita*), which involved a word-medial onset, as having the relatively greatest duration. Within this condition there was a further split determined by speech rate, with increased fricative duration in normal rate as compared to fast. Similarly, speech rate was responsible for the second split within the remaining conditions. However, the algorithm did not converge at that point, producing further splits. Conditions 4 and 7 clustered together within fast and normal speech rate. Both of these conditions involved word-final sibilants followed by a consonant (sonorant in Condition 4, voiced stop in Condition 7). Word-final pre-consonantal sibilants were relatively shorter than word-medial sibilants (Condition 2: *entusiasmo*, Condition 5: *esbozo*, Condition 6: *obispo*) as well as sibilants in derived onsets (Condition 3: *gas acre*). There was also an effect of speaker sex within Conditions 2, 3, 5 and 6 pronounced at fast rate, with male speakers showing increased fricative duration compared to female speakers.

3.2.1.4 *Intensity difference*

Figure 7 shows a tree for intensity difference measured in dB. This variable represents the absolute value of the difference between mean intensity of the

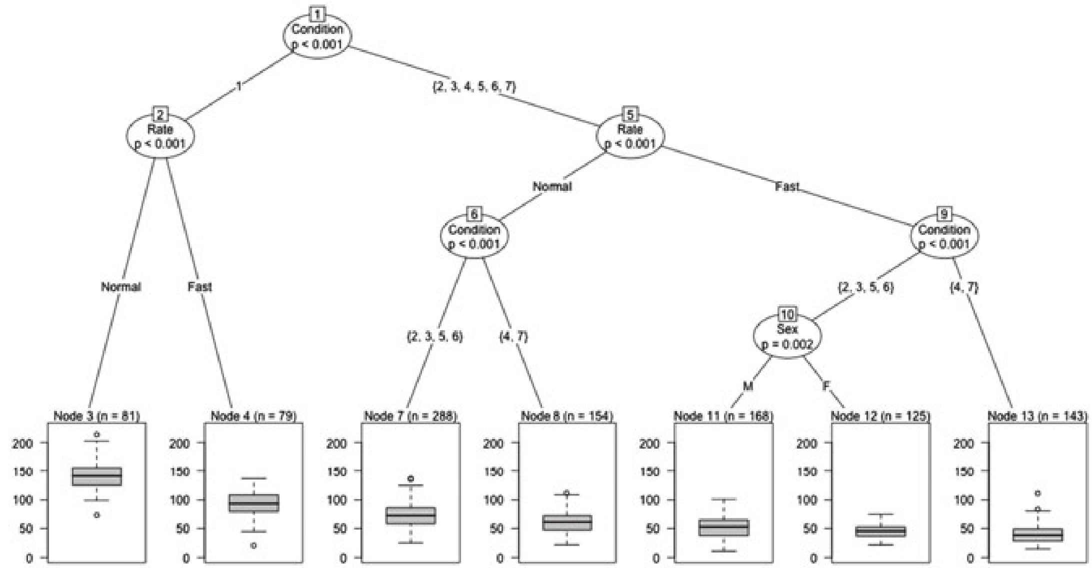

Figure 6
Conditional inference tree for fricative duration (in ms) based on a model with three predictors: condition, speech rate and speaker sex.

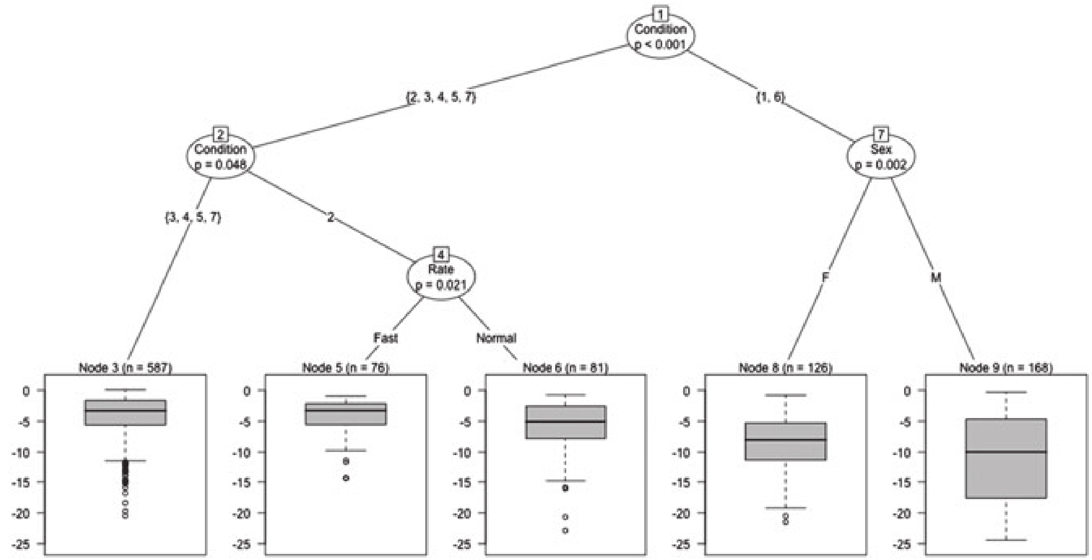

Figure 7
Conditional inference tree for intensity difference (in dB) based on a model with three predictors: condition, speech rate and speaker sex.

fricative and the mean intensity of its filtered low-frequency portion. Smaller differences are expected in fricatives with a high intensity low portion, as associated with voicing. This prediction is confirmed by the first split in Figure 7, which indicates that word-medial pre-vocalic sibilants (Condition 1: *casita*) and sibilants followed by voiceless stops (Condition 6: *obispo*) had a relatively larger intensity difference compared to all other sibilants. Within these two conditions there was a significant split conditioned by speakers sex, with male speakers showing a greater intensity difference than female speakers. Within the remaining conditions there was a split which singled out word-medial pre-sonorant sibilants (Condition 2: *entusiasmo*). Within this condition there was a difference between normal and fast speech condition, with a smaller intensity difference in the former case. Within the remaining conditions there was no effect of speech rate.

3.2.1.5 *Interim summary*

The partitioning analysis reveals a number of important generalisations concerning voicing in Quito Spanish sibilants. First, the partitioning confirms that the /s/-voicing environments involve increased voicing duration and voicing ratio when compared to the baseline involving word-medial pre-vocalic sibilants, or sibilants followed by voiceless stops. At the same time, however, the /s/-voicing contexts did not pattern consistently with sibilants followed by voiced stops. The two groups showed similar effects on voicing duration and voicing ratio, but only within fast speech rate. In the normal speech rate condition, sibilants followed by voiced stops showed higher voicing ratio and greater voicing duration than sibilants in the /s/-voicing contexts. The effect of speech rate confirms that /s/-voicing is sensitive to speech rate manipulations. A comparison of node 3 with nodes 5 and 6 in Figure 5 suggests that the voicing ratio increases for the /s/-voicing cases with an increase in speech rate. Furthermore, an effect of speaker sex was found within the /s/-voicing conditions at normal speech rate, with male speakers showing a greater voicing duration and voicing ratio than female speakers.

Condition and speech rate patterned somewhat differently in the models of fricative duration and intensity difference. Fricative duration appears to be mostly conditioned by prosodic factors, as the successive splits separate word-medial onsets from the remaining cases, and further reveal the clustering of word-final coda sibilants. In addition, the model of fricative duration confirms that fricatives were relatively longer in normal speech rate, as compared to fast. The intensity measure confirmed a difference between sibilants where no voicing is expected (word-medial pre-vocalic and pre-voiceless stops) from the remaining cases, but it did not produce a cluster of the /s/-voicing cases, and it was largely insensitive to speech rate manipulations. There were some significant effects of speaker sex in the models of fricative duration and intensity difference, but these effects were limited to specific conditions and speech rates.

3.2.2 *Speech rate effects and individual variation*

The results considered thus far invite a gradient interpretation of /s/-voicing. The voicing duration and ratio were shorter in the case of /s/-voicing than when a voiced stop followed within normal speech rate. In addition, the

/s/-voicing items had a higher voicing ratio in fast speech, which is consistent with a scenario where the same voicing duration coupled with fricative shortening conditioned by faster speech leads to an increase in voicing ratio. As a result, the /s/-voicing cases pattern with sibilants followed by voiced stops in fast speech. The significant effect of speaker sex within the /s/-voicing cases produced at normal rate also suggests a gradient interpretation of /s/-voicing, based on the prediction formulated in Section 1.4 above. The fact that speaker sex effects are only found in the /s/-voicing contexts, but not elsewhere, and only in normal speech rate, further reinforces the argument. Assuming that /s/-voicing is gradient, but voice assimilation to obstruents is categorical, only the former is expected to show sex effects. As far as speech rate is concerned, speaker sex effects may not be observable in fast speech rate, which is characterised by decreased fricative duration. If the duration of the constriction is sufficiently short, the intraoral pressure may not build up to the extent at which it quenches voicing, masking the effect of speaker sex. However, before we can conclude that /s/-voicing is gradient, we need to consider whether the effect of speech rate observed so far for the population also holds for every individual participant in the study, and whether the effect of speaker sex remains significant in a model that takes individual variation into account. In order to assess the effect of inter-speaker differences in the current data, we performed a mixed-effects regression analysis on the /s/-voicing data.

The analysis of individual variation was run on a subset of the data involving the /s/-voicing environments ($n = 428$) only. The partitioning analysis presented in Section 3.2.1 above confirms that /s/-voicing shows unique speech rate effects compared to other contexts tested in the study, and so subsetting the data for analysis of speech rate was done to avoid confounds from non-/s/-voicing environments. A series of models were fitted to the /s/-voicing data, using the lme4 package (Bates & Maechler 2009). The dependent variables in the subsequent models were: voicing duration, voicing ratio, fricative duration and intensity difference. The fixed predictors included in the models were the condition (word-medial pre-sonorant, word-final pre-vocalic, word-final pre-sonorant), speech rate (fast vs. normal) and speaker sex (male vs. female). All the fixed predictors were retained in the final models to allow a comparison with the analysis based on conditional inference trees presented in Section 3.2.1. We further verified, using the log-likelihood test, whether the inclusion of interactions between all or the subset of the predictors improved the fit of the model. If no significant improvement was found, the interactions were not retained. All the models had random intercepts for speaker and item, allowing for speaker-specific and item-specific variation from the overall mean. In addition, we considered the inclusion of a random slope for speech rate within speakers. For every dependent variable a model with a random slope for rate within speaker was compared to a model with a random-intercept only model in a log-likelihood

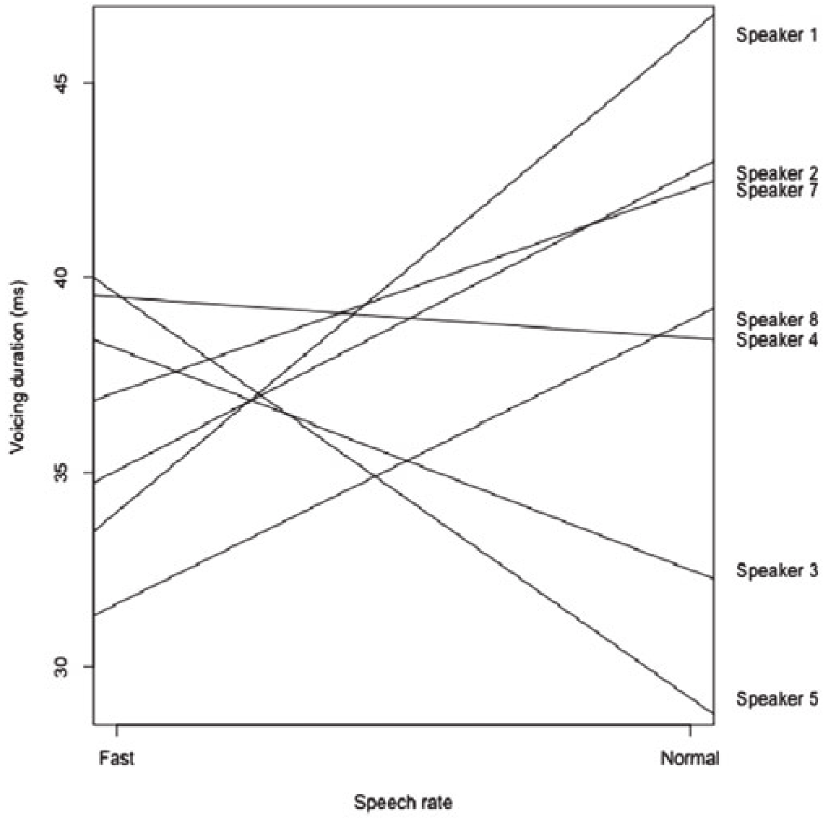

Figure 8
Coefficients of rate effects by speaker in a model with voicing duration (in ms) as a response variable, rate, condition and sex as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

test. In cases where random slopes were found to improve the fit they were retained in the model. We further considered whether the fit would improve further with an inclusion of a random effect of condition within speaker and random effect of rate and speaker sex within item.

3.2.2.1 *Voicing duration*

For voicing duration, a model with a random effect of rate within speaker was found to provide a significantly better fit than a model with a random intercept only, when compared in a log-likelihood test ($\chi^2 = 10.15$, $df = 2$, $p = .006$). The improvement in fit with an inclusion of a random slope confirms that speakers varied not only with respect to voicing duration in the fast rate, but also with respect to the adjustment they made in voicing duration when switching between normal and fast speech. The effects for individual speakers are presented in Figure 8. Two opposing tendencies can be observed in the population. For speakers 3, 4 and 5 the duration of voicing increased at fast speech rate, in line with the previous result based on conditional inference trees. This effect was largest for speaker 5, and very limited for speaker 4. Speakers 1, 2, 7 and 8, on the other hand, produced more

Term	Level	β	*SE*	*t*
Intercept		36.41	2.93	12.43
Rate	Normal	2.19	3.96	0.55
Condition	2 (*entusiasmo*)	−3.36	2.39	−1.41
Condition	4 (*gas noble*)	−7.27	2.44	−2.98
Sex	M	5.78	2.86	2.02

Table 3
Regression coefficients, with standard error, and *t* values for a model predicting the duration of voicing (in ms) in the /s/-voicing environments. The intercept corresponds to a word-final /s/ before a vowel (e.g. *gas acre* 'acrid gas') pronounced by a female speaker at a fast rate.

voicing at normal as opposed to fast speech rate. The effect was greatest for speaker 1, and smallest for speaker 7. For speakers 3, 4 and 5 the duration of voicing increased at fast speech rate. This effect was largest for speaker 5, and very limited for speaker 4.

No further improvements in the model fit were obtained by adding random slopes for condition within speaker and rate within word. Thus, only the random slope for rate within speaker was retained in the final model. Table 3 presents a summary of the fixed effects.[7] Voicing duration was greater in word-final pre-vocalic sibilants (*gas acre*) than word-medially before a sonorant (*entusiasmo*), although the relatively low *t*-value ($t = -1.41$) suggests that the difference between these two conditions might not be significant. The duration of voicing in word-final pre-vocalic sibilants was also greater than in word-final sibilants followed by a sonorant consonant. The *t*-value associated with this difference was considerably greater ($t = -2.98$). A log-likelihood comparison of the model in Table 3 with a nested model without the main effect of condition showed a significant difference ($\chi^2 = 7.84$, $df = 2$, $p = .02$), consistent with an interpretation that condition is a significant predictor. There was no significant effect of speech rate for the population, given that the value of the standard error exceeded the value of the β coefficient. This finding is consistent with the large amount of individual variation found in the model; once the individual opposing tendencies are taken into account, no significant generalisation can be made for the population. Male speakers showed increased voicing duration compared to female speakers. However, a log-likelihood comparison of the

[7] Because we are not aware of a reliable method for estimating p-values in models with random slopes, p-values are not included in the model summary for this or subsequent models. In our discussion of the models in this section, we assume that $|t| > 2$ might indicate a significant main effect. As this approach is somewhat informal, in crucial cases, we use the log-likelihood test, comparing nested models to assess whether individual predictors were significant. We report the relevant results throughout this section.

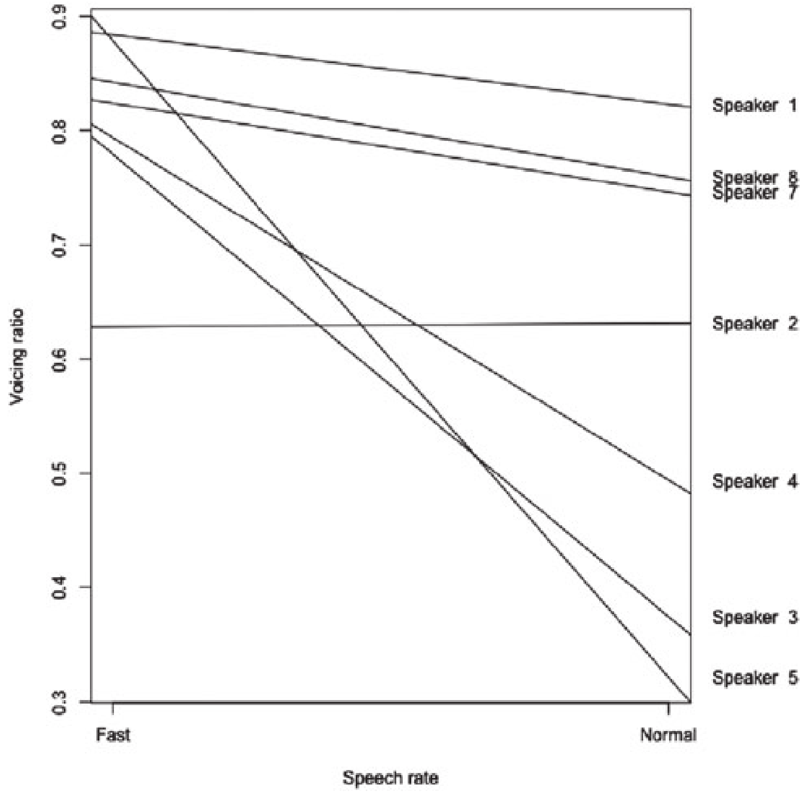

Figure 9
Coefficients of rate effects by speaker in a model with voicing ratio as a response variable, rate, condition and speaker sex as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

model in Table 3 with a nested model without the main effect of sex showed no significant difference in fit, ($\chi^2 = 2.37$, $df = 1$, $p = .11$), indicating that sex is not a significant predictor in modelling voicing duration.

3.2.2.2 *Voicing ratio*

Similarly to the case of voicing duration, the modelling of voicing ratio showed a significant improvement to the fit once speech rate was considered as a random effect within speaker ($\chi^2 = 29.51$, $df = 2$, $p < .001$). The two models compared had rate and condition as fixed effects, and random intercepts for speaker and item. Figure 9 presents the slopes and intercepts for individual speakers' voicing ratio. For speakers 3, 4 and 5 there was a large difference between normal and fast speech rate, with a higher voicing ratio in the latter case. Speakers 1, 7 and 8 also had a higher voicing ratio in fast speech rate, but for those speakers the adjustment between fast and normal speech rate was considerably lower, with a smaller difference in voicing ratio between fast and normal speech. Speaker 2 was atypical in producing (minimally) higher voicing ratio at normal speech rate as opposed to fast.

Term	Level	β	*SE*	*t*
Intercept		0.80	0.08	10.57
Rate	Normal	−0.21	0.09	−2.35
Condition	2 (*entusiasmo*)	−0.001	0.06	−0.01
Condition	4 (*gas noble*)	0.06	0.06	1.01
Sex	M	0.004	0.08	0.05

Table 4
Regression coefficients, with standard error, and *t* values for a model predicting the voicing ratio in the /s/-voicing environments. The intercept corresponds to a word-final /s/ before a vowel (e.g. *gas acre* 'acrid gas') pronounced by a female speaker at a fast rate.

No further improvements in the model fit were obtained by adding random slopes for condition within speaker and rate within word. Table 4 presents the fixed effects structure of the model with fixed effects of rate and condition, and random effects of speaker, word and rate within speaker. There was a negative effect of rate, with higher voicing ratio at fast rate compared to normal. The effects of condition and speaker sex on voicing ratio were very small, as indicated by small β- and t-values. The fit of the model was not significantly reduced upon removing condition ($\chi^2 = 1.37$, $df = 2$, $p = .5$) or speaker sex as a predictor ($\chi^2 = 0$, $df = 1$, $p = 1$). Unfortunately, the residuals of this model are non-normally distributed, as in a large number of cases the voicing ratio equalled 1, and so the results of the model in Table 4 must be approached with caution.

3.2.2.3 *Fricative duration*

Similarly to the case of voicing duration and ratio, individual effects were considered in mixed-effects modelling of fricative duration for the /s/-voicing data. The response variable has a skewed distribution and initial modelling of the fricative duration returned a series of models with non-normally distributed residuals. To remedy this problem, a Box-Cox transformation (Box & Cox 1964) was applied to this response variable. Based on the optimal λ-value of .5, the square root of the fricative duration was subsequently analysed, with an improvement to the normality of the response variable and the distribution of the residuals.

Just like in previous cases, inclusion of a random slope for rate within speaker significantly improved the fit of the model compared to a model with random intercepts only ($\chi^2 = 24.43$, $df = 2$, $p < .001$). However, no further significant improvements were found with the inclusion of random slopes for rate within item, or condition within speaker. The results of the model comparison show that there were significant individual differences with respect to how fricative duration was affected by speech rate manipulation.

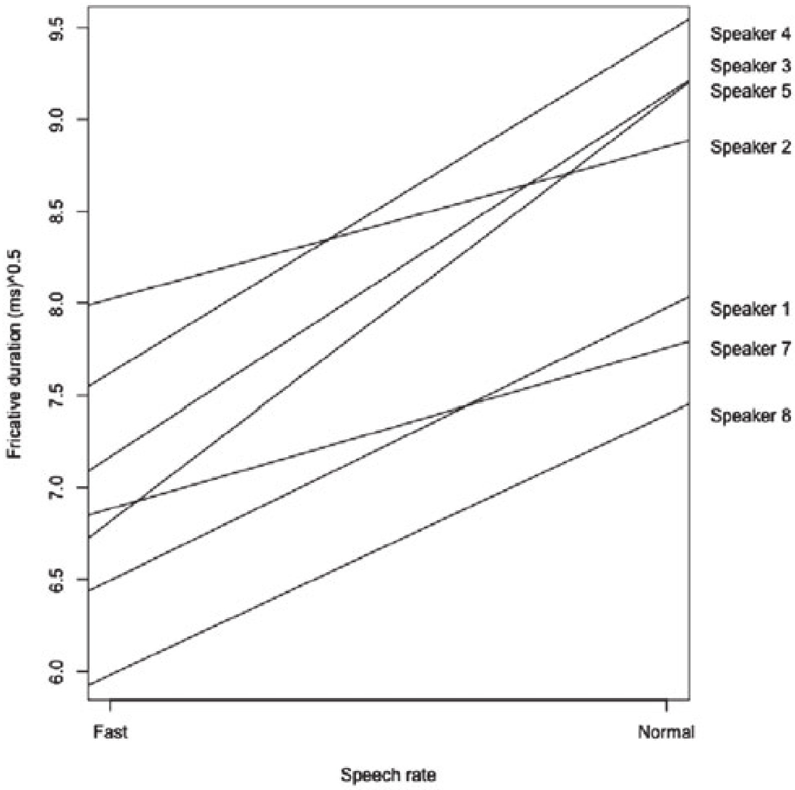

Figure 10

Coefficients of rate effects by speaker in a model with square root of fricative duration (in ms) as a response variable, rate, condition and sex as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

A plot of slopes and intercepts for individual speakers is in Figure 10. All the participants had longer fricative duration at normal speech rate as compared with fast, which is expected. However, speakers 3, 4 and 5 had both higher intercepts and higher slopes than speakers 1, 7 and 8. Speaker 2 had a relatively high intercept, but a shallow slope, again patterning differently from the remaining speakers.

Table 5 shows a summary of the fixed effects in the model. Fricative duration was greatest in Condition 3 (*gas acre*). The difference in duration was small compared to Condition 2 (*entusiasmo*, $\beta = -0.25$, $SE = 0.28$, $t = -0.91$), but quite considerable compared to Condition 4 (*gas noble*, $\beta = -0.97$, $SE = 0.28$, $t = -3.45$). Condition was a significant predictor based on a log-likelihood comparison of the model in Table 5 with a nested model where the fixed effect of condition was not included ($\chi^2 = 10.60$, $df = 2$, $p = .005$). There was a relatively large effect of speech rate, with an increase in fricative duration at normal speech rate ($\beta = 1.53$, $SE = 0.24$, $t = 6.41$). Sex of the speaker had a very small effect on fricative duration ($\beta = 0.15$, $SE = 0.58$, $t = 0.26$), and it was not found to improve the fit of the model, based on the log-likelihood test ($\chi^2 = 0$, $df = 1$, $p = 1$).

Term	Level	β	*SE*	*t*
Intercept		7.00	0.48	14.50
Rate	Normal	1.53	0.29	6.41
Condition	2 (*entusiasmo*)	−0.25	0.28	−0.91
Condition	4 (*gas noble*)	−0.97	0.28	−3.45
Sex	M	0.15	0.58	0.26

Table 5
Regression coefficients, with standard error, and *t* values for a model predicting the square root of fricative duration (in ms) in the /s/-voicing environments. The intercept corresponds to a word-final /s/ before a vowel (e.g. *gas acre* 'acrid gas') pronounced by a female speaker at a fast rate.

3.2.2.4 *Intensity difference*

Similarly to the case of fricative duration, a Box-Cox transformation was applied to intensity difference in order to obtain a normal distribution. The absolute value of the intensity difference was raised to the power of 0.22. Model comparison revealed that there were significant individual differences with respect to the effect of speech rate. A model with fixed effects of condition and rate, random intercepts for speaker and item and random slopes for rate within speaker was found to have a better fit than a corresponding model without random within-speaker effects ($\chi^2 = 19.97$, $df = 2$, $p < .001$). No significant improvement in the fit of the model was found with an inclusion of random effects of rate within item, or condition within speakers. Figure 11 illustrates the effect of speech rate manipulation on intensity difference for individual speakers. Speakers 3, 4 and 5 had steep positive slopes, with larger difference between low-frequency intensity and overall mean intensity in normal speech rate, compared to fast. Speakers 7 and 8 also had positive, but more shallow slopes, with relatively less adjustment in intensity difference between speech rates. Speakers 1 and 2 showed the opposite tendency, with greater intensity difference in normal speech, compared to fast, but the difference between speech rates within these two speakers was smaller than within speakers 3, 4 and 5.

Table 6 shows the coefficients for the fixed part of the model. There was very little variation associated with condition, and removing condition as a predictor from the model did not significantly affect the model's fit ($\chi^2 = 2.01$, $df = 2$, $p = .37$). The same was found for the effect of speaker sex, as a model without sex as a fixed predictor did not differ significantly in its fit from the model presented in Table 6 ($\chi^2 = 0$, $df = 1$, $p = .01$). There was a positive main effect of rate ($\beta = 0.07$, $SE = 0.04$, $t = 1.73$) although, as previously discussed, this effect varied depending on the individual speaker.

Term	Level	β	*SE*	*t*
Intercept		1.26	0.08	16.27
Rate	Normal	0.07	0.04	1.73
Condition	2 (*entusiasmo*)	0.03	0.03	1.05
Condition	4 (*gas noble*)	−0.01	0.03	−0.26
Sex	M	0.01	0.1	0.12

Table 6

Regression coefficients, with standard error, and *t* values for a model predicting intensity difference between mean low-frequency intensity and total mean intensity (in dB) in the /s/-voicing environments (absolute value raised to the power of 0.22). The intercept corresponds to a word-final /s/ before a vowel (e.g. *gas acre* 'acrid gas') pronounced by a female speaker at a fast rate.

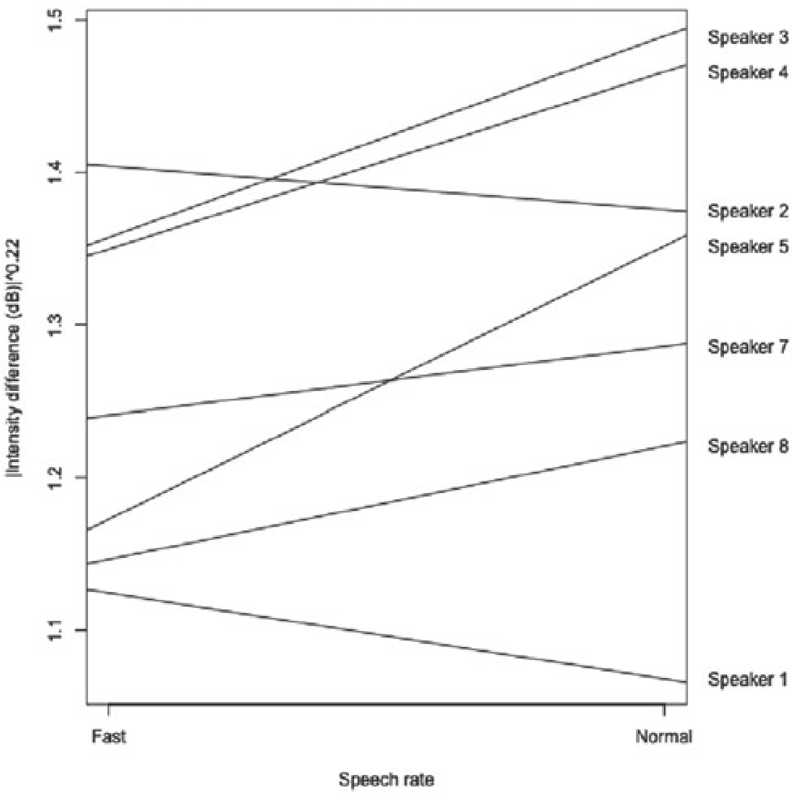

Figure 11

Coefficients of rate effects by speaker in a model with intensity difference (in dB) raised to the power of 0.22 as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

3.2.3 *Analysis of speech rate*

Since durational measurements, including voicing duration and fricative duration, are influenced by speech rate, an analysis of speech rate was performed to help contextualise previous findings concerning duration, with a view to determining which durational adjustments are likely to be an effect of

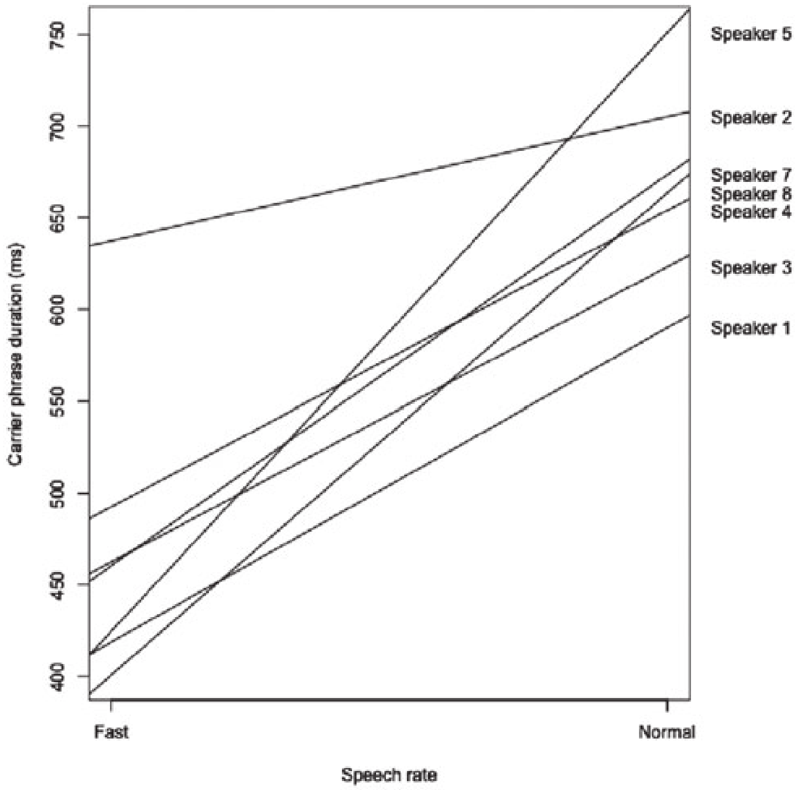

Figure 12
Coefficients of rate effects by speaker in a model with the duration of the phrase *otra vez* "one more time" (in ms) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

individual speech rate, as opposed to individual voicing strategy. Speech rate was quantified as the duration of the phrase *otra vez* measured in ms. A linear mixed effects model with the duration of the phrase *otra vez* as a dependent variable was fitted to the data from all the environments (n = 1038). The effects of speaker and item were treated as random, and rate was included as a fixed effect, and as a random effect within speaker. This model was found in a log-likelihood test to provide a significantly better fit of the data in comparison with a model where the rate was included as a fixed effect only ($\chi^2 = 128.18$, $df = 2$, $p < .001$). This shows that individual speakers differed not only in the intercept, but also in how much faster their speech was when they were trying to speak fast, as compared to the speech rate they found comfortable. The model was further trimmed to remove 18 outliers with a standardised residual greater than 2.5 standard deviations from 0, based on the observation of skewness in the distribution of the residuals.

Figure 12 illustrates speech rate coefficients for individual speakers in the trimmed model. In the fast speech rate condition Speaker 2 was by far the slowest one, using over 100 ms more than others to produce the phrase *otra vez*. This speaker also showed the smallest difference between the two speech rates. Speaker 5 showed the relatively greatest difference between fast and

Term	Level	β	*SE*	*t*
Intercept		471.15	30.31	15.54
Rate	Normal	194.72	31.68	6.15

Table 7
Regression coefficients, with standard error, and *t* values for a model predicting the duration of the phrase *otra vez* 'one more time' (in ms). The intercept corresponds to the fast rate.

normal speech rate, being relatively slowest when speaking normally, but patterning with others in the fast rate condition. The remaining speakers showed some differences with respect to slope and intercept, but no distinctive individual strategies emerge from the data. Table 7 presents the summary of the fixed part of the model. As expected, rate had a major effect of on the duration of the *otra vez*. The duration of the phrase was on average 194.72 ms greater at the normal speech rate compared to fast ($SE = 31.68$, $t = 6.15$).

3.2.4 *Summary*

Our findings confirm that sibilants in the /s/-voicing conditions are subject to increased voicing compared to sibilants followed by voiceless stops or sibilants in word-medial onsets. In a partitioning analysis, the /s/-voicing conditions showed increased voicing duration and ratio and smaller intensity difference, compared to the voiceless baseline. At the same time, the /s/-voicing conditions did not pattern consistently with cases where a sibilant was followed by a voiced obstruent, taken here to represent the voiced baseline. Specifically, the /s/-voicing cases showed relatively smaller voicing duration and ratio within normal speech rate. In fast speech, on the other hand, the two groups patterned together, typically with full voicing during frication.

Based on the partitioning analysis, there were no differences between contexts reported as environments for Quito /s/-voicing. Word-medial and word-final pre-sonorant /s/ and word-final pre-vocalic /s/ patterned together in models of voicing duration, voicing ratio and intensity difference. Further analysis of the effect of condition using mixed-effects models showed that word-final pre-vocalic sibilants (Condition 3, *gas acre*) had greater voicing duration than word-final pre-sonorant sibilants (Condition 4, *gas noble*). However, in a model of fricative duration, sibilants in Condition 3 were also found to be longer than those in Condition 4. Consistent with these findings, the two groups showed no differences in models of voicing ratio or intensity difference, and condition was not found to significantly improve the fit of either of those models.

The effect of speech rate on voicing duration and ratio within the /s/-voicing contexts uncovered by the partitioning analysis could be taken as an indication of a gradient status of Quito Spanish /s/-voicing. This generalisation is supported by the fact that voicing ratio was higher in fast speech than in normal speech in the /s/-voicing contexts, whereas there was a limited effect of rate on voicing duration. However, a closer analysis of the rate effects on Quito voicing reveals that this generalisation does not apply to all the participants in the study. While some speakers behaved consistently with the direction of the effect found by the partitioning analysis, others did the opposite.

Models of voicing duration, voicing ratio, fricative duration and intensity difference all showed a significantly improved fit once speech rate was considered as a random effect within speaker. This confirms that individual speakers responded differently to speech rate manipulations. Four out of seven participants in the experiment (speakers 1, 2, 7 and 8) showed increased voicing duration in normal as compared with fast speech rate within the /s/-voicing environments. The remaining three speakers (3, 4 and 5) showed the opposite tendency, producing more voicing in the fast speech rate, as compared with normal. Those speakers who produced more voicing in the fast rate condition also had a greater difference in the duration of fricatives produced at normal and fast rate. For speakers 1, 2, 7 and 8 the difference was relatively smaller, even though there was also an increase in the duration of a fricative from fast to normal speech condition. Finally, distinct individual strategies emerge from the data with respect to how rate affected voicing ratio and the difference between mean low- and total-frequency intensity in fricatives. For speakers 3, 4 and 5 a large difference was observed between normal and fast rate in models of both voicing ratio and intensity difference. For these speakers an increase in speech rate brought about an increase in voicing ratio, and the value of intensity difference was closer to zero. Speakers 1, 2, 7 and 8, on the other hand, maintained a similar voicing ratio and intensity difference across varying speech rates.

The two groups of speakers which emerge from the analysis of the random effect of rate were not consistently differentiated by sex. Speakers 1, 2 and 8 were male, but speaker 7 was female. Within the other group, speakers 3 and 5 were female, but speaker 4 was male. A main effect of speaker sex was found in the partitioning models: male speakers had higher voicing duration and ratio, compared to female speakers, within the /s/-voicing cases in normal speech. However, the main effect of sex was not significant in any of the mixed models. The effect of sex approached significance in the model of voicing duration (as indicated by the *t*-value of 2.02), but removing this effect did not significantly affect the models fit. In the models of voicing ratio and intensity difference, the effect of sex was small, as were the associated *t*-values, and including the effect of sex did not significantly improve the fit. These results indicate that the inter-speaker variation found in the data is more adequately explained as individual variation rather than as variation conditioned by speaker sex.

Finally, an analysis of speech rate was performed in order to verify whether any previously observed durational trends were also reflected in the duration of the phrase *otra vez*. The analysis showed that speaker 2 had a relatively small adjustment in the duration of the phrase between normal and fast speech rate. Speaker 5 showed a relatively large difference in the duration of the carrier phrase between normal and fast speech rates, consistent with the previously noted large adjustment in fricative duration between the two conditions. However, speakers 3 and 4, who had patterned with 5 with respect to large difference in fricative duration, did not also show the same effect with respect to overall speech rate, having a relatively small difference in the duration of the phrase *otra vez* between normal and fast speech rate. Unlike in the case of voicing duration and fricative duration, no two distinct trends emerged from the measurements of speech rate.

4. Discussion

4.1 */s/-voicing: Categorical or gradient?*

The interpretation of the current data that we shall pursue is that Quito /s/-voicing is optional but categorical for some speakers, and gradient for others. All speakers showed variation in their realisation of /s/-voicing in producing a varying duration of vocal fold vibration during the fricative. At first blush, the intra-speaker variation could be taken as supporting the hypothesis that /s/-voicing is gradient, i.e. involves varying degrees of laryngeal overlap between the fricative and the neighbouring segment. However, some aspects of the variation that we find in the data do not straightforwardly support the gradient interpretation for all the speakers.

An argument against the hypothesis that /s/-voicing in Quito Spanish is gradient for all speakers comes from speech rate effects. Drawing on previous work by Solé (1992, 1995), we predicted that a gradient process of /s/-voicing should involve either no variation in voicing duration across different speech rates, or potentially increased voicing at higher speech rates due to increased gestural overlap. The productions by speakers 3, 4 and 5 fit this description. Speaker 4 produced a similar duration of voicing across different speech rates, while for speakers 3 and 5 the duration of voicing increased somewhat in the fast speech condition. This effect, coupled with fricative duration shortening at fast speech rate, led to a significant increase in the voicing ratio from normal to fast speech rate for all three speakers. The increase in ratio was reflected in the increase in intensity difference between normal and fast speech rate. Partially voiced fricatives are expected to have relatively higher intensity difference than fully voiced ones, as the presence of voiceless frication will both reduce the low-frequency intensity and increase the mean total intensity. In that way, intensity difference can be said to be negatively correlated with voicing ratio.

For categorical voicing processes we expected to see no variation in ratio across different speech rates, coupled with increased voicing duration at normal speech rate. This was observed for four speakers: 1, 2, 7 and 8. Speaker 2 was somewhat exceptional in showing speech rate effects consistent with a categorical voicing hypothesis, while mostly producing partially voiced tokens. However, evidence from speech rate effects has to be taken with caution in case of speaker 2, as the speech rate analysis shows an exceptionally small difference in the carrier phrase duration, which raises the question whether the speaker really completed the speech rate manipulation task. Speakers 1, 7 and 8 showed a clear preference for fully voiced fricatives regardless of speech rate, and they differed from speakers 3, 4 and 5 in producing the majority of fricatives as fully voiced at normal speech rate. The effects of a stable voicing ratio across speech rates was reflected for speakers 1, 7 and 8 in intensity difference, which was similar in fast and normal speech for these speakers.

There are no phonetic reasons why gradient coarticulation should involve more voicing at slower speech rates. However, an increase in the voicing duration coupled with an increase in duration of the fricative makes sense as a strategy to realise a specific phonetic target. Therefore, the realisation of /s/-voicing by speakers 1, 7 and 8 is consistent with that of an optional phonological rule, which applies only to a subset of the /s/-voicing cases. When it does apply, however, the rule is categorical and reflected in full phonetic voicing.

The split into two groups of speakers as proposed above based on the analysis of rate effects is not readily explained by speaker sex. One of the hypotheses we formulated in Section 1.4 specified that the occurrence of speaker sex related effects in Quito /s/-voicing could indicate a gradient nature of the process. The partitioning analysis showed that within the /s/-voicing contexts, male speakers had a greater voicing duration than female speakers. However, this effect was not found to be significant in a mixed-effect model. These findings are not necessarily contradictory. Out of the three speakers whom we analyse as gradient (3, 4 and 5), two were female and one was male, and out of the three categorical speakers (1, 7 and 8), two were male and one was female. Categorical speakers are expected to show increased mean voicing duration compared to gradient speakers. In addition to this, male gradient speakers are expected to show increased voicing duration compared to female gradient speakers. Both of these predictions are reflected in our data (see voicing duration values for individual speaker at normal speech rate in Figure 8 above). These two factors taken together are likely to produce a significant effect of speaker sex in a model which does not take individual variation into account, hence the significant speaker sex effect in the partitioning analysis. However, a closer analysis reveals a more complex pattern of variation in the data. Mixed-effects models show that there is a split in the population reflecting how speakers are affected by rate. The individual strategies with respect to rate manipulations explain the variation better than a fixed effect of speaker sex, making the effect of sex not significant.

Another argument for identifying two distinct voicing strategies in the population of speakers studied in this paper comes from the results of fricative duration modelling. Speakers 1, 7 and 8 patterned together with respect to fricative duration, producing shorter fricatives than speakers 3, 4 and 5. The difference between the two groups is especially visible at normal speech rate (see Figure 10 above). The analysis of the carrier phrase duration shows that speakers 1, 7 and 8 did not overall speak slower than speakers 3, 4 and 5. Therefore, the relative decrease in fricative duration in normal speech rate is most likely a part of a voicing strategy employed by speakers 1, 7 and 8, as decreased fricative duration is a potential cue to voicing. The fact that speech rate manipulations reveal a difference in the population of speakers with respect to fricative duration is also relevant to the discussion of whether the difference between the two groups can be attributed to differences in speaker sex. Unlike voicing duration, fricative duration is not expected to be directly affected by the length of the vocal tract, so we would not expect this variable to be affected by sex. Therefore, the observed differences in fricative duration between speakers 1, 7 and 8, and 3, 4 and 5 reinforce the argument that two distinct voicing strategies can be discerned for the experiment participants that are independent of their sex.

The analysis proposed here concerning categoricity and gradience in /s/-voicing extends to all potential /s/-voicing environments, i.e. coda sibilants followed by a sonorant consonant in the same word (*entusiasmo*) or in the next word (*gas noble*), and word-final sibilants followed by a vowel in the next word (*gas acre*). We did not find significant differences in voicing ratio or intensity difference between these three environments. While pre-vocalic /s/-voicing showed increased voicing duration compared to word-medial pre-sonorant context, a similar difference was also found between these two conditions with respect to fricative duration. The effect of increased voicing in the pre-vocalic environment is consistent with a voicing strategy which we propose for the categorical speakers, and which involves adjusting different voicing cues (in this case, voicing duration) to maintain a stable voicing ratio in the /s/-voicing cases.

4.2 *Theoretical implications*

The categorical status of /s/-voicing, even if optional and subject to individual variation, presents a serious challenge to Colina's (2009) analysis, which, as previously discussed, crucially relies on /s/-voicing being gradient. Colina (2009) models /s/-voicing as a strictly phonetic effect operating on top of abstract phonology, which requires word-final pre-pausal and word-final pre-vocalic sibilants to have identical representations in terms of phonological features. However, if word-final pre-vocalic sibilants are optionally voiced at the phonological level, as follows from our data, their phonological form cannot be said to be identical to the sibilants in the word-final pre-pausal context where voicing does not apply.

Importantly, the significance of the Quito Spanish data goes beyond challenging a single analysis, as it shows the limits of modelling opaque patterns as surface correspondence effects. The optional but categorical voicing of /s/ in cases like [ga.za.kɾe] 'acrid gas' cannot be straightforwardly relegated from the domain of phonology. However, it cannot be analysed as a morphophonological correspondence effect either, as there is no principled means of selecting a base form that could serve as a source of Output–output voicing correspondence. The word-final sibilant in *gas* surfaces as voiced only in phrasal contexts, such as *gas noble* or *gas blanco*, and there is no morphosyntactic motivation for why these forms should be considered the base rather than other phrasal contexts with no /s/-voicing, such as utterance-final *gas*.

The patterns of variation found in our data also question the analysis by Bradley (2005) and Bradley & Delforge (2006), which is based on the maintenance of systemic contrasts. This analysis proposes that coda /s/ is underspecified for voicing in the output of lexical phonology. In post-lexical phonology, the underspecification is maintained for pre-consonantal /s/. Pre-vocalic /s/, on the other hand, is proposed to undergo categorical voicing, driven by a systemic requirement to maintain maximal possible contrast. This requirement is, in turn, linked to the observation that pre-vocalic /s/-voicing, but not pre-consonantal /s/-voicing, creates phrasal minimal pairs, as exemplified in (8). Therefore, the analysis crucially distinguishes between gradient pre-consonantal voicing and categorical pre-vocalic voicing. This distinction does not find confirmation in our data, which show that /s/-voicing can be either categorical or gradient, but that the variation depends on the individual speaker rather than on the following segment. We find both gradient voicing of word-final /s/ before a vowel, and categorical coda /s/-voicing before a sonorant consonant. From the formal point of view, the analysis by Bradley (2005) and Bradley & Delforge (2006) could perhaps be extended to accommodate gradient pre-vocalic voicing and categorical pre-consonantal voicing. However, in the absence of an empirical link between the occurrence of contrast and either categoricity or gradience in the realisation of /s/-voicing, the explanation based on systemic contrast maintenance and the emergence of phonological /s/-voicing is a lot less compelling.

In comparison, none of the findings of our study poses a problem to the cyclical analysis. Phrase-level /s/-voicing may be optional, and gradient effects follow from a modular view of the phonetics–phonology interface,

Stage	Change	Examples
1	Initial stage	[gas] [gas.no.βle] [ga.sa.kɾe]
2	Coda delaryngealisation	[gaS] [gaS.no.βle] [ga.sa.kɾe]
3	Analogical change	[gaS] [gaS.no.βle] [ga.Sa.kɾe]
4	Categorical reinterpretation	[gaS] [gaz.no.βle] [ga.za.kɾe]

Table 8
The life cycle of /s/-voicing.

where phonetics operates on the output of phonological computation.[8] The fundamental insight of the cyclical analysis is that /s/-voicing applies to those sibilants that had been in a coda at some stage of the derivation. While the derivational character of this generalisation makes it incompatible with monostratal models, it can in fact emerge from a series of natural and well attested sound changes. Table 8 presents a succession of phonetic and phonological changes through which /s/-voicing can evolve, drawing on the insights and generalisations of the life cycle of a phonological process (Bermúdez-Otero 2007, Bermúdez-Otero & Trousdale 2012). At stage 1, the voicing contrast is lost in sibilants by way of devoicing. Early Spanish had a complex inventory of sibilant phonemes, consisting of several pairs of voiced and voiceless members, but by the sixteenth century various developments had reduced this inventory to a single sibilant: /s/. At stage 2, syllable codas are interpreted as laryngeally underspecified (see discussion of causes for coda delaryngealisation below). The pre-vocalic voicing at stage 3 in Table 8 is where the cyclic architecture becomes crucial, as this change is phonological rather than phonetic and proceeds by input restructuring. In cyclic terms, the process of coda delaryngealisation climbs up from the Phrase Level to the Word Level, as the learners reinterpret the output of the Phrase Level (in our case coda delaryngealisation), as being already present in the input, i.e. at the output of the Word Level. With delaryngealisation applying at Word Level, all word-final fricatives lose their voice specifications, becoming amenable to Phrase Level voicing before consonants and vowels alike. This type of voicing originates as gradient, but later undergoes stabilisation, represented as stage 4 in Table 8.

Pre-sonorant voicing of syllable- and word-final sibilants might have its roots in aerodynamic characteristics of coda fricatives. Solé (2010) compares

[8] Optionality of post-lexical rules is, in fact, widely observed. Examples include variable realisation of English r-sandhi (Giegerich 1999), English /n#k/ sandhi, where some speakers vary between categorical assimilation and no assimilation (Ellis & Hardcastle 2002), as well as Italian nasal place assimilation, which tends to be categorical, but which may be suspended when there is a word boundary intervening, particularly in slower speech (Celata et al. 2013).

coda and onset fricatives based on simultaneous acoustic and aerodynamic data, and shows that coda fricatives show a reduced oral gesture, a lower oral pressure build-up, lower velocity of air through the oral constriction and a less intense frication. According to Solé (2010), delayed onset of frication is of consequence to language change, as it may result in a shorter fricative, which is more likely to get affected by overlapping gestures in a pre-consonantal environment. Gradient voicing resulting from increased overlap might, with time, become categorical through perceptual reinterpretation on the part of listeners. This type of reinterpretation is consistent with experimental findings which show that increasing the portion of voicing in a fricative can lead to categorically voiced perceptions (Forrez 1966, Stevens et al. 1992). Alternatively, a similar perceptual effect can also follow directly from the decreased duration and intensity of frication in coda fricatives, as both of these acoustic cues are associated with fricative voicing (Slis & Cohen 1969a, b; Stevens et al. 1992; Ladefoged & Maddieson 1996). Duration-sensitive perceptual reinterpretation is also an alternative to the gestural overlap explanation for why voicing takes place pre-consonantally, but not pre-pausally, as the pre-pausal environment is associated with lengthening (Klatt 1975, Cooper & Danly 1981) which potentially hinders voiced percepts.

As far as word-final pre-vocalic fricatives are concerned, their voicing does not follow directly from the phonetic conditioning discussed thus far. Due to resyllabification, word-final pre-vocalic /s/ is not expected to display the aerodynamic characteristics of coda fricatives, which is a prerequisite to the occurrence of passive voicing in the previously discussed scenario. The voicing effect also cannot be ascribed to segmental factors alone; if /s/-voicing were triggered by mere adjacency of sibilants to phonetically voiced sonorants and vowels, we would expect voicing to apply also in word-medial onsets. The fact that in present-day Quito Spanish /s/-voicing applies in the word-final pre-vocalic and pre-sonorant context alike suggests that the pre-vocalic voicing is a result of analogical change. The idea is that language users extend an already existing process of word-final pre-sonorant voicing to a novel context, i.e. word-final pre-vocalic /s/. This kind of change is reminiscent of the spread of /t/-glottalisation from pre-consonantal to the pre-vocalic environment in Southern British English (Williams & Kerswill 1999). Rácz (2011) argues that /t/-glottalisation is phonetically motivated before a consonant, but not before a vowel. Thus, pre-vocalic glottalisation emerges not due to phonetic conditioning, but as an analogical extension from the pre-consonantal environment. Assuming a rich memory model, in which phonetic detail is stored together with individual productions of lexical entries, and a categorisation model in which a learner classifies tokens as belonging to a specific category based on their similarity to other tokens, Rácz (2011) models a learner who begins glottalisating before a vowel due to the influence of pre-consonantal /t/-glottalisation and 10% production noise, where pre-vocalic tokens are erroneously coded by the learner as pre-consonantal and vice versa.

Similarly, as in British English /t/-glottalisation, Quito /s/-voicing is not phonetically motivated in word-final position before a vowel. However, pre-vocalic voicing might emerge by means of analogical change, as proposed by Rácz (2011) for British English, or by input restructuring, as proposed within the cyclic account by Bermúdez-Otero (2011). The cyclic account also relies on analogy, albeit conceptualised differently than in the frequency-based model, as at the historical stage preceding the stabilisation of voicing, de-laryngealisation is levelled from pre-pausal and pre-consonantal tokens to pre-vocalic ones, with the effect that the process percolates from the Phrase Level to the Word Level. Interestingly, this analogical relationship between word-final pre-vocalic and word-final pre-sonorant /s/ is incompatible with the assumptions of Output–output correspondence. While the analogical relationship between *gas noble* and *gas acre* can potentially arise from factors such as phonetic similarity or frequency, neither of these form a part of selection criteria for what makes an appropriate base for transderivational correspondence.

A *JL* referee points out that the diachronic scenario goes against the proposal by Robinson (1979, 2012) that word-final pre-vocalic /s/-voicing in Quito Spanish is not an innovation, but a remnant of the /s/–/z/ contrast in Medieval Spanish. Robinson (1979, 2012) argues that the sibilant voicing contrast was present in the early colonisation period in Ecuador, due to the influence of settlement from Andalusia, where the contrast was retained for a relatively long time. The loss of sibilant voicing contrast was a later development in both Andalusia and Ecuador. According to Robinson, this later influence took a long time to spread to Ecuadorian highlands, and it is likely that some remnants of the contrast survive in the form of pre-vocalic voicing. Bradley & Delforge (2006) reject Robinson's diachronic assumption, arguing that the lack of lexical voicing contrast in Ecuadorian Spanish indicates that this dialect had gone through an earlier stage of post-lexical devoicing. We would like to add that the diachronic development of /s/-voicing, as proposed by Robinson, is unexpected from the point of view of functional factors driving sound change development. It is a very unusual situation where a language loses a lexical contrast, but simultaneously keeps the same phonetic contrast at the post-lexical level, where the functional pressure to keep different forms apart is relatively lower. What is more, our findings on /s/-voicing in present-day Quito Spanish show that pre-vocalic /s/-voicing (where post-lexical voicing contrast is created) patterns consistently with pre-sonorant coda /s/-voicing (where no contrast is involved). This finding is consistent with an analysis where voicing follows from voicing neutralisation in codas at the lexical level, but it does not relate in any obvious way to the potential contrastive function of /s/-voicing, or to the way the contrast functioned in medieval Spanish.

The life cycle of /s/-voicing, as outlined in Table 8 above, also predicts the kind of variation we find in Quito Spanish, where a categorical and a gradient process coexist within the same speech community. The gradient version of /s/-voicing is schematised in stage 3. It applies following the analogical change for the delaryngealisation rule (which makes word-final pre-vocalic voicing pattern with the pre-sonorant cases), but before the voicing rule has been re-analysed as categorical. Sibilants in the /s/-voicing context are delaryngealised in the output of the phonological grammar at this stage, to later undergo gradient voicing in the phonetics. The coexistence of a more conservative gradient pattern with an innovative categorical variant is not an unusual situation from the point of view of the life cycle. Rather, it is a natural transition characteristic of sound change in progress.[9] Bermúdez-Otero (2011, 2013) uses the term 'rule scattering' to refer to cases where a gradient and categorical rules operate simultaneously in a language, and identifies instances of it in e.g. Philadelphia /æ/-tensing (Labov 1989, 2010) and English /l/-darkening (Sproat & Fujimura 1993, Boersma & Hayes 2001, Yuan & Liberman 2009). The rule scattering analysis can also be extended to the case of English /n#k/ sandhi, which Ellis & Hardcastle (2002) found to apply gradiently or categorically, depending on the speaker (recall Section 1.4 above). Ellis & Hardcastle (2002) stress that marked individual differences in external sandhi patterns may be obscured in analyses that average over speakers. Our own findings also support this important methodological conclusion, showing that what appears to be a gradient pattern in an analysis of a population might in fact comprise a mixture of categorical and gradient variation. Another consequence of this is that rule scattering may be quite common, but underrepresented in linguistic literature, as it requires both instrumental evidence and analysis of individual variation in order to be identified.

5. Conclusion

The acoustic data collected in the current study show that the phonetic production of Quito /s/-voicing involves categorical behaviour on the part of some speakers, whereas other speakers show characteristics of phonetic gradience. We have argued for an interpretation of these facts consistent with a modular view of the phonetics–phonology interface, and we consider

[9] A related issue here is that of the phonetics–phonology interface. A *JL* referee notes that coexistence of categorical and gradient versions of the same process within a speech community potentially complicates a rigid phonetics–phonology distinction which we assume in our analysis. This is an important point, which deserves more detailed consideration than we are able to give within the confines of this paper. For a discussion on how the life cycle relates to the issue of modularity in grammar, the reader is referred to Chapter 2 in Strycharczuk (2012) and to Bermúdez-Otero (2013).

the categorical behaviour to signify an optional assignment of a phonological category. Rejecting the hypothesis that all /s/-voicing is necessarily gradient is of crucial consequence to a correspondence-based analysis of phonological opacity; although there are forms where the realisation of word-final /s/ as voiced is transparent, there is no principled way of selecting those forms as the base for the opaque word-final pre-vocalic sibilants. The case of synchronic phonological opacity we find in Quito Spanish appears to have originated from analogy between lexical forms which are not in a transderivational correspondence relationship. As a consequence of the analogical change, two steps of derivation seem necessary for a formal model to capture the resulting synchronic pattern, if the goal of the model is to provide a common generalisation for all the environments where /s/-voicing applies.

APPENDIX A

Items used in the experiment

Condition 1. Word-internal pre-vocalic	
gasita	'gauze (DIM)'
casita	'home (DIM)'
aviso	'notice'
gusano	'worm'
mayonesa	'mayonaisse'
mosaico	'mosaic'
Condition 2. Word-internal pre-nasal	
entusiasmo	'enthusiasm'
espasmo	'spasm'
budismo	'buddhism'
bautismo	'baptism'
esmoquín	'the tuxedo'
cosmólogo	'cosmologist'
Condition 3. Word-final pre-vocalic	
gas acre	'acrid gas'
palmas altas	'tall palm trees'
tesis obvia	'obvious thesis'
virus asnal	'brutal virus'
tres autores	'three authors'
muchos hombres	'many men'
Condition 4. Word-final pre-nasal	
gas noble	'noble gas'
ropas negras	'black clothes'
croquis nuevo	'new foundation'
crisis mundial	'world crisis'
bases nuevas	'naval bases'
muchos monjes	'many monks'

Condition 5. Word-internal before a voiced stop

esbozo	'plan, sketch'
Lisboa	'Lisbon'
jurisdicción	'jurisdiction'
presbítero	'priest'
posdata	'postscript'

Condition 6. Word-internal before a voiceless stop

subasta	'auction'
carraspera	'transport'
obispo	'bishop'
crepúsculo	'dawn'
asbesto	'asbesthos'
microscopio	'microscope'

Condition 7. Word-final before a voiced stop

marchas buenas	'good marches'
velas blancas	'white sails'
brindis digno	'dignified toast (fig.)'
cactus grande	'big cactus'
grandes barcos	'big ships'
gatos bellos	'beautiful cats'

APPENDIX B

Means and standard deviations for the four dependent variables analysed in the current study by condition, rate and sex

Voicing duration (ms)					
Condition	Rate	Sex	N	Mean	SD
1. *gasita*	Fast	F	34	9.30	10.90
		M	45	8.61	10.54
	Normal	F	34	6.14	8.32
		M	47	7.06	8.11
2. *entusiasmo*	Fast	F	31	38.53	12.73
		M	45	35.90	19.02
	Normal	F	35	26.80	19.12
		M	46	46.84	26.71
3. *gas acre*	Fast	F	33	37.11	15.54
		M	43	41.49	21.21
	Normal	F	25	32.71	23.10
		M	28	51.69	22.96
4. *gas noble*	Fast	F	27	30.64	10.58
		M	41	33.44	16.61
	Normal	F	33	32.55	17.40
		M	41	36.88	21.42

5. *esbozo*	Fast	F	33	35.82	14.97
		M	44	34.88	17.33
	Normal	F	36	46.92	23.63
		M	48	53.00	27.74
6. *obispo*	Fast	F	28	9.03	6.20
		M	36	11.17	8.88
	Normal	F	30	7.39	7.27
		M	40	10.66	6.65
7. *marchas buenas*	Fast	F	33	34.38	13.65
		M	42	31.65	16.51
	Normal	F	35	47.98	24.11
		M	45	51.10	25.09

Voicing ratio					
Condition	Rate	Sex	N	Mean	SD
1. *gasita*	Fast	F	34	0.12	0.17
		M	45	0.12	0.23
	Normal	F	34	0.04	0.06
		M	47	0.05	0.05
2. *entusiasmo*	Fast	F	31	0.90	0.24
		M	45	0.77	0.35
	Normal	F	35	0.43	0.36
		M	46	0.69	0.37
3. *gas acre*	Fast	F	33	0.79	0.33
		M	43	0.79	0.33
	Normal	F	25	0.51	0.40
		M	28	0.82	0.34
4. *gas noble*	Fast	F	27	0.85	0.28
		M	41	0.84	0.30
	Normal	F	33	0.61	0.37
		M	41	0.72	0.38
5. *esbozo*	Fast	F	33	0.84	0.30
		M	44	0.78	0.35
	Normal	F	36	0.77	0.34
		M	48	0.74	0.36
6. *obispo*	Fast	F	28	0.22	0.20
		M	36	0.20	0.14
	Normal	F	30	0.09	0.09
		M	40	0.15	0.11
7. *marchas buenas*	Fast	F	33	0.94	0.21
		M	42	0.79	0.34
	Normal	F	35	0.78	0.35
		M	45	0.84	0.32

Fricative duration (ms)

Condition	Rate	Sex	N	Mean	SD
1. *gasita*	Fast	F	34	86.45	16.50
		M	45	97.42	27.09
	Normal	F	34	136.48	16.91
		M	47	146.91	29.68
2. *entusiasmo*	Fast	F	31	44.34	11.08
		M	45	49.21	14.78
	Normal	F	35	72.77	22.61
		M	46	72.76	20.01
3. *gas acre*	Fast	F	33	49.01	11.77
		M	43	54.71	17.20
	Normal	F	25	69.19	13.32
		M	28	66.12	13.85
4. *gas noble*	Fast	F	27	39.34	14.51
		M	41	42.69	17.42
	Normal	F	33	60.92	19.04
		M	41	56.42	22.20
5. *esbozo*	Fast	F	33	43.80	11.57
		M	44	49.70	17.86
	Normal	F	36	63.29	20.06
		M	48	75.18	22.21
6. *obispo*	Fast	F	28	45.13	12.85
		M	36	54.32	19.93
	Normal	F	30	81.24	16.68
		M	40	75.80	21.31
7. *marchas buenas*	Fast	F	33	36.29	11.18
		M	42	44.78	18.29
	Normal	F	35	63.52	15.54
		M	45	62.60	18.73

Intensity difference (dB)

Condition	Rate	Sex	N	Mean	SD
1. *gasita*	Fast	F	34	8.30	3.84
		M	45	11.27	7.72
	Normal	F	34	9.14	4.78
		M	47	12.53	7.46
2. *entusiasmo*	Fast	F	31	3.68	2.15
		M	45	4.65	3.38
	Normal	F	35	6.37	3.60
		M	46	5.58	4.99

3. *gas acre*	Fast	F	33	4.41	3.21
		M	43	4.15	3.08
	Normal	F	25	6.29	3.84
		M	28	2.78	1.82
4. *gas noble*	Fast	F	27	3.60	1.63
		M	41	4.40	3.10
	Normal	F	33	5.24	3.33
		M	41	4.44	4.35
5. *esbozo*	Fast	F	33	4.12	2.99
		M	44	4.09	3.44
	Normal	F	36	4.33	2.49
		M	48	4.29	4.64
6. *obispo*	Fast	F	28	7.19	4.78
		M	36	9.13	5.78
	Normal	F	30	9.54	4.85
		M	40	11.20	7.40
7. *marchas buenas*	Fast	F	33	3.09	2.84
		M	42	4.15	3.75
	Normal	F	35	4.05	3.60
		M	45	3.60	3.70

REFERENCES

Alvord, Scott M. (ed.). 2011. *5th Conference on Laboratory Approaches to Romance Phonology*. Somerville, MA: Cascadilla Proceedings Project.

Barry, Martin C. 1992. Palatalisation, assimilation and gestural weakening in connected speech. *Speech Communication* 11, 393–400.

Bates, Douglas & Martin Maechler. 2009. *lme4: Linear mixed-effects models using S4 classes*. http://CRAN.R-project.org/package=lme4, accessed 13 August 2010. R package version 0.999375-32.

Bermúdez-Otero, Ricardo. 2007. Diachronic phonology. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 497–517. Cambridge: Cambridge University Press.

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In van Oostendorp et al. (eds.), 2019–2048.

Bermúdez-Otero, Ricardo. 2013. Amphichronic explanation and the life cycle of phonological processes. In Patrick Honeybone & Joseph C. Salmons (eds.), *The Oxford handbook of historical phonology*. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo & Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *Handbook on the history of English: Rethinking and extending approaches and methods*, 691–720. New York: Oxford University Press.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32, 45–86.

Boersma, Paul & David Weenink. 2009. Praat: Doing phonetics by computer (Version 5.1.12). http://www.praat.org/, accessed 15 October 2009.

Box, George E. P. & David R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society* (Series B (Methodological)) 26, 211–252.

Bradley, Travis. 2005. Sibilant voicing in Highland Ecuadorian Spanish. *Lingua* (*gem.*) 2, 29–42.

Bradley, Travis & Ann Marie Delforge. 2006. Systemic contrast and the diachrony of Spanish sibilant voicing. In Randall Gess & Deborah Arteaga (eds.), *Historical Romance linguistics: Retrospectives and perspectives*, 19–52. Amsterdam: John Benjamins.

Celata, Chiara, Silvia Calamai, Irene Ricci & Chiara Bertini. 2013. Nasal place assimilation between phonetics and phonology: An EPG study of Italian nasal-to-velar clusters. *Journal of Phonetics* 41, 88–100.

Chappell, Whitney. 2011. The intervocalic voicing of /s/ in Ecuadorian Spanish. In *5th Workshop on Spanish Sociolinguistics*, 57–64. Somerville, MA: Cascadilla Proceedings Project.

Cohn, Abigail C. 1993. Nasalization in English: Phonology or phonetics. *Phonology* 10, 43–82.

Cohn, Abigail C. 2006. Is there gradient phonology? In Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds.), *Gradience in grammar: Generative perspective*, 25–44. Oxford: Oxford University Press.

Colina, Sonia. 2009. Sibilant voicing in Ecuadorian Spanish. *Studies in Hispanic and Lusophone Linguistics* 2, 3–29.

Connell, Bruca & Amalia Arvaniti (eds.). 1995. *Phonology and phonetic evidence: Papers in Laboratory Phonology IV*, 315–333. Cambridge: Cambridge University Press.

Cooper, William E. & Martha Danly. 1981. Segmental and temporal aspects of utterance-final lengthening. *Phonetica* 38, 106–115.

Crystal, Thomas H. & Arthur S. House. 1988. A note on the durations of fricatives in American English. *Journal of the Acoustical Society of America* 84, 1932–1935.

Cuartero Torres, Néstor. 2001. *Voicing assimilation in Catalan and English*. Ph.D. dissertation, Universitat Autònoma de Barcelona.

Ellis, Lucy & William J. Hardcastle. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics* 30, 373–396.

Ernestus, Mirjam. 2011. Gradience and categoricality in phonological theory. In van Oostendorp et al. (eds.), 2115–2136.

Ernestus, Mirjam & [R.] Harald Baayen. 2006. The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In Louis M. Goldstein, Douglas Whalen & Catherine T. Best (eds.), *Laboratory Phonology 8: Varieties of phonological competence* (Phonology and Phonetics 4), 27–49. Berlin & New York: Mouton De Gruyter.

Flemming, Edward Stanton. 1995. *Auditory representations in phonology*. Ph.D. dissertation, University of California, Los Angeles.

Flemming, Edward Stanton. 2002. *Auditory representations in phonology*. New York: Routledge.

Forrez, Gabriël. 1966. *Relevante parameters van de stemhebbende fricatief /z/*. Eindhoven: Inst. Perceptie Onderzoek Verslag.

García, Christina. 2011. Intervocalic /s/ voicing in the Spanish of Loja, Ecuador. Master's thesis, The Ohio State University.

Giegerich, Heinz J. 1999. *Lexical strata in English: Morphological causes, phonological effects* (Cambridge Studies in Linguistics 89). Cambridge: Cambridge University Press.

Gradoville, Micheal Stepehen. 2011. Validity in measurements of fricative voicing: Evidence from Argentine Spanish. In Alvord (ed.), 59–74.

Harris, James W. 1969. *Spanish phonology*. Cambridge, MA: MIT Press.

Harris, James W. 1983. *Syllable structure and stress in Spanish: A nonlinear analysis*. Cambridge, MA: MIT Press.

Holst, Tara & Francis Nolan. 1995. The influence of syntactic structure on [s] to [ʃ] assimilation. In Connell & Arvaniti (eds.), 315–333.

Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15, 651–674.

Hualde, José Ignacio. 2005. *The sounds of Spanish*. Cambridge: Cambridge University Press.

Jansen, Wouter. 2004. *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch*. Ph.D. dissertation, University of Groningen.

Jessen, Michael. 2009. Instability in the production and perception of intervocalic closure voicing as a cue to *bdg* vs. *ptk* in German. *Folia Linguistica* 38, 27–42.

Keating, Patricia A. 1988. Underspecification in phonetics. *Phonology* 5, 275–292.

Keating, Patricia A. 1990. Phonetic representations in a generative grammar. *Journal of Phonetics* 18, 321–334.
Keating, Patricia A. 1996. The phonology–phonetics interface. In Ursula Kleinhenz (ed.), *Interfaces in phonology* (Studia Grammatica 41), 262–278. Berlin: Akademie.
Kiparsky, Paul. 1982. From Cyclic Phonology to Lexical Phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations*, vol. 1, 131–175. Dordrecht: Foris.
Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2, 83–138.
Klatt, Dennis H. 1975. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3, 129–140.
Labov, William. 1989. The child as linguistic historian. *Language Variation and Change* 1, 85–97.
Labov, William. 2010. *Principles of linguistic change: Cognitive and cultural factors*. Malden, MA: Wiley-Blackwell.
Ladefoged, Peter. 1973. Features of the larynx. *Journal of Phonetics* 1, 73–83.
Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Malden, MA: Blackwell.
Lipski, John M. 1989. /s/-voicing in Ecuadoran Spanish: Patterns and principles of consonantal modification. *Lingua* 79, 49–71.
Myers, Scott. 2000. Boundary disputes: The distinction between phonetic and phonological sound patterns. In Noel Burton-Roberts, Philip Carr & Gerard J. Docherty (eds.), *Phonological knowledge*, 245–272. Oxford: Oxford University Press.
Nolan, Francis, Tara Holst & Barbara Kühnert. 1996. Modelling [s] to [ʃ] accommodation in English. *Journal of Phonetics* 24, 113–137.
Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.
Pierrehumbert, Janet B. 2006. The next toolkit. *Journal of Phonetics* 34, 516–530.
Prince, Alan & Paul Smolensky. 2004 [1993]. *Constraint interaction in generative grammar*. Malden, MA & Oxford: Blackwell.
R Development Core Team. 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org.
Rácz, Péter. 2011. To err is human – rich memory language models and t-glottalisation in English. Presented at Old World Conference in Phonology (OCP8).
Robinson, Kimball L. 1979. On the voicing of intervocalic *s* in the Ecuadorian Highlands. *Romance Philology* 33, 132–143.
Robinson, Kimball L. 2012. The dialectology of syllabification: A review of variation in the Ecuadorian Highlands. *Romance Philology* 66, 115–145.
Ryan, Kevin. Forthcoming. Contextual and non-contextual prosodic minimality. In Lena Fainleib, Nicholas LaCara & Yangsook Park (eds.), *41st Annual Meeting of the North East Linguistic Society* (NELS 41). Amherst, MA: GLSA, University of Massachusetts.
Schmidt, Lauren B. & Erik W. Willis. 2011. Systematic investigation of voicing assimilation of Spanish /s/ in Mexico City. In Alvord (ed.), 1–20.
Slis, Iman Hans & Antonie Cohen. 1969a. On the complex regulating the voiced–voiceless distinction I. *Language and Speech* 12, 80–102.
Slis, Iman Hans & Antonie Cohen. 1969b. On the complex regulating the voiced–voiceless distinction II. *Language and Speech* 12, 137–155.
Solé, Maria-Josep. 1992. Phonetic and phonological processes: The case of nasalization. *Language and Speech* 35, 29–43.
Solé, Maria-Josep. 1995. Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. *Language and Speech* 38, 1–23.
Solé, Maria-Josep. 2010. Effects of syllable position on sound change: An aerodynamic study of final fricative weakening. *Journal of Phonetics* 38, 289–305.
Sproat, Richard & Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21, 291–311.
Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge, MA: MIT Press.

Stevens, Kenneth N., Sheila E. Blumstein, Laura Glicksman, Martha Burton & Kathleen Kurowski. 1992. Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America* 91, 2979–3000.

Strycharczuk, Patrycja. 2012. *Phonetics–phonology interactions in pre-sonorant voicing*. Ph.D. dissertation, University of Manchester.

Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24, 135–178.

Tucker, Benjamin V. & Natasha Warner. 2010. What it means to be phonetic or phonological: The case of Romanian devoiced nasals. *Phonology* 27, 289–324.

van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.). 2011. *Blackwell companion to phonology*. Malden, MA: Wiley-Blackwell.

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

Williams, Ann & Paul Kerswill. 1999. Dialect levelling: Continuity vs. change in Milton Keynes, Reading and Hull. In Paul Foulkes & Gerard J. Docherty (eds.), *Urban voices: Accent studies in the British Isles*, 141–162. London: Arnold.

Williams, Lee. 1977. The voicing contrast in Spanish. *Journal of Phonetics* 5, 169–184.

Yuan, Jiahong & Mark Liberman. 2009. Investigating /l/ variation in English through forced alignment. In *INTERSPEECH2009: 10th Annual Conference of the International Speech Communication Association*, 2215–2218. Bristol: International Speech Communication Association.

Zsiga, Elizabeth C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In Connell & Arvaniti (eds.), 282–302.

Authors' addresses: *(Strycharczuk)*
School of Languages, Linguistics and Culture,
The University of Manchester, Oxford Road,
Manchester M13 9PL, UK
pstrycharczuk@gmail.com

(Van 't Veer)
Leiden University Centre for Linguistics, Leiden University,
Postbus 9515, 2300 RA Leiden, The Netherlands
B.M.van.t.Veer@hum.leidenuniv.nl

(Bruil)
Leiden University Centre for Linguistics, Leiden University,
Postbus 9515, 2300 RA Leiden, The Netherlands
m.bruil@hum.leidenuniv.nl

(Linke)
Leiden University Centre for Linguistics, Leiden University,
Postbus 9515, 2300 RA Leiden, The Netherlands
& The Meertens Institute, PO Box 94264, 1090 GG Amsterdam,
The Netherlands
kathrin.linke@meertens.knaw.nl