Sound initiation and source types in human imitations of sounds

Pétur Helgason
Department of Linguistics & Philology, Uppsala University, Sweden, and
Speech, Music and Hearing, KTH, Sweden
petur.helgason@lingfil.uu.se

Abstract

There exists a rich body of research exploring the production of speech, but for non-linguistic sound production, for example imitations of environmental sounds or animals, much less data and research are available. Data from human sound imitations collected in the initial, exploratory phase of the SkAT-VG project were analyzed in terms of the articulatory and aerodynamic conditions involved in their production. These exploratory data yielded a classification of sound productions in imitations based on the intersections between sound initiation and sound source types. The source types identified are turbulent, myoelastic, whistled and percussive sources. The ways in which these source types intersect with pulmonic, glottalic and velaric sound initiation, both egressive and ingressive, are described and discussed.

Introduction

In speech, the principal way of producing sound is to drive an airstream past one or more obstacles. The organ responsible for driving the airstream is the initiator (Pike 1943: 85ff), while the source of the sound produced is located at the point of the obstacle(s).

The sound initiation mechanisms commonly acknowledged in speech production are pulmonic egressive, glottalic egressive, glottalic ingressive and velaric ingressive (ibid.; see also Catford, 1977). Although there are no attested cases of pulmonic ingressive and velaric egressive airstreams being utilized as features in phonological systems, there is no real obstacle to producing sounds using these initiation mechanisms. Pulmonic ingressive sounds, in particular, are quite common (cf. Eklund, 2008), and also occur in imitations. Sounds can also be produced without creating an airstream, e.g. by clashing the teeth together or by slapping the tongue against the floor of the mouth. Such sounds are referred to as percussives (Pike, 1943: 103). Percussives are encountered in sound imitations, but they are rarely found in (non-pathological) speech.

The source-filter model of speech production (Fant, 1960) has been successful in describing the acoustics of human speech sound production. In speech the principal sound sources are voicing, produced with a pulmonic egressive airstream entraining the vocal folds into vibration, and friction noise, produced by constricting a pulmonic egressive airstream at some point in the vocal tract, causing turbulence. However, humans can produce sounds with a number of additional source types, some of which are used in spoken languages and some of which are not.

Sound initiation and source types

Here, the focus is on cataloguing source types that seem useful for sound imitation. The approach is to categorize the source types according to the articulatory and aerodynamic conditions under which they are produced. The main categories of source types thus identified are **myoelastic**, **turbulent**, **whistled** and **percussive**. The three former source types can be produced using various initiation mechanisms, but percussives constitute an initiation mechanism on their own. In the following,

examples of these four basic types of sources will be discussed primarily in terms of the initiation mechanisms involved and their observed or potential uses in sound imitation.

The exploratory data have various sources. Many of the examples on which the analyses are based have been found on-line, but exploratory recordings have also been made, with the aid of a professional improvisational actor.

Turbulent sources

To produce fricative sounds, an airstream is made turbulent by channeling it through a constriction in the glottis or the vocal tract (cf. Stevens 1999: 37f for an overview). In the exploratory phase of the SkAT-VG project we have observed imitations using pulmonic egressive friction (which parallel fricatives in speech) as well as velaric ingressive friction (which parallel clicks or click-like sounds). We have no examples yet where imitators use pulmonic ingressive, glottalic egressive or ingressive or velaric egressive friction.

Pulmonic egressive turbulence

Friction made with a pulmonic egressive airstream is by far the most commonly occurring turbulent source in imitations, just as it is in speech. As is the case with speech sounds, a turbulent friction noise can be made at many places in the vocal tract. This type of friction is especially common in the imitation of "basic" sound events, such as the interaction of solids (e.g. knocking, scraping and squeaking sounds) and sounds of gases in motion (e.g. blowing, puffing and hissing sounds) (cf. Lemaitre et al. 2011 for further examples of sound events). For example, the impression given by an improvisational actor of the sound of "scraping on a hard surface" is quite speech-like and can be described as a voiceless velar fricative [x].

Pulmonic ingressive turbulence

While pulmonic ingressive friction is not difficult to produce, it is difficult (or impossible) to produce sibilant fricatives with an ingressive airstream (Catford, 1988: 20ff; see also Eklund 2008 for a more comprehensive review). In other cases, although appreciably different, the acoustic result of ingressive friction is still quite similar acoustically to the egressive counterpart. These facts may contribute to its apparent scarcity in imitations. However, one should note that ingressive friction is encountered in emotive sounds, e.g. sucking in air through one's teeth to indicate pain (Cruttenden 1986: 180).

Glottalic egressive turbulence

Glottalic egressive friction is fairly common in languages, but as yet unattested in our exploratory data of imitations. Possibly, the acoustically similar outcomes of glottalic and pulmonic egressive friction are a contributing factor – why use a glottalic airstream when a pulmonic airstream creates, more or less, the same sound?

Glottalic ingressive turbulence

According to UPSID (Maddieson and Precoda, 1990) voiceless glottalic ingressive speech sounds (i.e., voiceless implosives) are phonologically distinctive in less than 1% of the world's languages. Judging by this typological rarity one could assume that such sounds are fairly difficult to produce. The exploratory data have not yet yielded imitations that make use of a glottalic ingressive airstream, as such. However, note Pike's (1943: 40) observation that English speakers sometimes use a voiceless velar implosive [ɠ] to imitate the "glug-glug" sound of pouring liquid from a bottle (the voiced counterpart can also be used). Thus, despite the typological rarity of such sounds, they still seem to be used in imitations.

Velaric egressive turbulence

A velaric egressive source has not been encountered in the exploratory data, but one can conceive of such sounds being used to imitate sputtering in liquids.

Squeezing a velaric airstream out between the teeth, for example, may faithfully replicate the sound of a spraying can (although, obviously, this depends on denture). An ingressive airstream leads to an acoustically similar result.

Velaric ingressive turbulence

Velaric ingressive turbulence is used to produce click sounds, which are typologically rare. Still, paralinguistic click sounds are encountered quite frequently in speech (cf., e.g., Jakobson, 1979: 40). In English, for example, the dental click even has a more or less standardized orthography, variably written as *tut-tut* or *tsk-tsk*.

In the SkAT-VG exploratory data set, the impression of "trickling water" made by an improvisational actor contains an example of velaric ingressive initiation (see Figure 1). To achieve this effect, the actor alternated soft postalveolar or alveolar click sounds with sublaminal percussives (discussed below, in the section on percussives) with frequent and rapid labial modifications of the resonance characteristics.

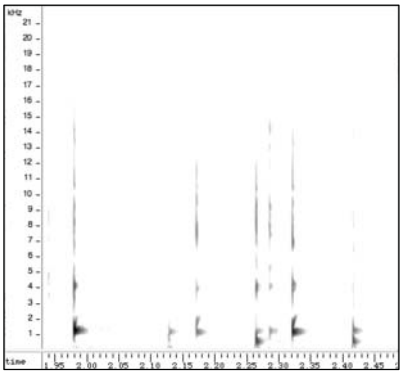

Figure 1. A spectrogram of an actor's impression of the sound of "trickling water".

Myoelastic sources

In the myoelastic source type, muscle and elastic tissue are made to oscillate in an air stream. This can lead to (almost) periodic sounds or intermittent breaks in an otherwise turbulent airstream. Crucially, for some myoelastic source types the oscillation is frequent enough to be perceived as a tone.

Pulmonic egressive myoelastic sources

The most commonly encountered myoelastic source by far, both in speech and sound imitations, is pulmonic egressive vocal fold phonation, i.e., voicing. As a sound source in speech and singing, the vocal folds are highly versatile, allowing a great deal of precision in the control of onset and offset, timbre and oscillation frequency.

In linguistic phonetics, a distinction is made between several vocal fold phonation types. Modal voice, breathy voice and creaky voice are the principal types (stiff voice, slack voice are also recognized but are not considered here, nor is the difference between breathy voice and whispery voice; see Ladefoged & Maddieson (1996) for an overview of the linguistic uses of voice). Non-linguistic voicing types include falsetto and pressed voice.

These various voice qualities are relevant for sound imitations, perhaps most notably in the imitation of animal sounds and engine sounds. The imitation of a cow, for example, usually involves a modal voice quality with a nasal resonance. The croak of a frog may be imitated with a creaky voice quality (an ingressive creak works even better). Falsetto voice is frequently encountered in animal imitations, e.g. when imitating a cat meowing.

A much less common myoelastic source type is aryepiglottic phonation in which the aryepiglottic folds vibrate in an air stream at frequencies ranging from approximately 40 to 100 Hz (Moisik, Esling & Crevier-Buchman, 2010). In the exploratory data we have observed impressions of animal growling in which aryepiglottic phonation is used, but usually it is used in combination with voicing. Similarly, there are examples of imitations of rumbling engines, which combine aryepiglottic vibration and voicing.

At least four types of supralaryngeal pulmonic egressive myoelastic sources

can be created. First, some people can achieve a uvular myoelastic oscillation, equivalent to uttering a voiceless, uvular trill, [ʀ̥]. Second, some people can achieve an apico-alveolar oscillation, equivalent to producing a voiceless, apico-alveolar trill, [r̥]. For these two source types, the rate of oscillation can exceed 30 Hz, but they are still not perceived as tones but rather as a rapid series of impacts. There are no examples of these two source types being used on their own for imitations in the exploratory data, but there are examples of the apico-alveolar source combined with whistling in bird imitations.

A third supralaryngeal source type uses a dorso-lateral configuration for the tongue and pushes out air between the tongue dorsum and a stricture that appears to be located at or anterior to the palatoglossal arch. The sound produced is periodic with an f0 range from approximately 150 and 700 Hz, judging from the examples gathered so far. The most well known use of the dorso-lateral source type is the voice of Donald Duck, the famous cartoon character. The exploratory data set contains numerous examples of the use of this source type in the imitation of birds.

The fourth supralaryngeal source type is made with a bilabial constriction. The constriction can be made with two distinct lip configurations, which yield quite different results. First, the lips can be pressed together without much stiffness in the labial tissue while an airstream is passed through. This leads to a fairly slow periodic myoelastic vibration (25-35 Hz) that is not perceived as a tone. The exploratory data set contains an example of such a voiceless, bilabial trill being used to imitate the blowing sound of a horse. The second lip configuration involves pressing the lips together quite tightly and making them much stiffer while forcing an airstream between them. This can lead to a (multiply) periodic source, which, in the exploratory data set, is found in the imitation of an elephant trumpeting.

Pulmonic ingressive myoelastic sources

When vocal fold phonation is made with a pulmonic ingressive airstream the result is ingressive voicing. Acoustically, ingressive voicing is quite distinct from egressive voicing, sounding harsher and less sonorant (cf. Eklund, 2008). Like egressive voicing, ingressive voicing can be made both as ingressive falsetto and ingressive creak.

In imitations, an ingressive falsetto is quite common. It is used to imitate various animal sounds, such as a dog bark, a pig squeal and crow caw, but it can also be used to imitate squeaking sounds, such as the squeaking sound of wiping a window pane.

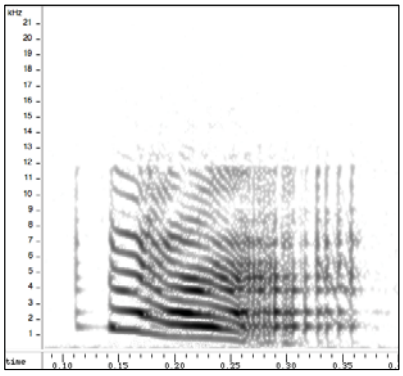

Figure 2. A spectrogram of an actor's impression of a "squeak from a window pane".

One example in the exploratory data set, shown in Figure 2, does contain both ingressive falsetto and ingressive creak. This is the impression made by an improvisational actor of the sound of a "squeak from a window pane".

Glottalic and velaric myoelastic sources

The SkAT-VG exploratory data set contains no imitations that make use of glottalic and velaric airstreams coupled with a myoelastic source. Using glottalic and velaric airstreams there is a very limited volume of air available to drive a myoelastic oscillation. Some configurations do yield a myoelastic effect, for example a glottalic egressive airstream can be coupled with an apico-alveolar

source to produce the equivalent of an ejective trill, [r̰']. However, the fact that these types of sources cannot be sustained for very long, if it can be achieved at all, reduces their usefulness in imitations, which may explain their absence in the exploratory data set.

Whistled sources

Very few languages are reported to have distinctive whistled coronal sibilants (Shosted, 2006). According to Shosted (ibid.: 566), whistled sibilants are produced in a manner similar to "a form of recreational whistling referred to as 'palatal' or 'roof' whistling", which is achieved by letting the tongue tip form a constriction that directs the airflow to the edges of the teeth. Pure "palatal" whistling is seldom encountered except in the repertoire of whistling virtuosi, such as the Hungarian Hacki Tamás or the Australian Luke Janssen. Still, the exploratory data set does include an example of this type of whistling being used to imitate the American Robin (*Turdus migratorius*).

In languages that do not have distinctive sibilant whistling, whistling can still occur sporadically when apical sibilants are produced and sibilants with a whistled component, similar to those found in speech, are observed when people imitate wind or weather noise.

Labial whistling does not occur in speech but the majority of people appear to be able to produce some form a labial whistle and this type of whistling is encountered frequently in daily life. Typically, labial whistling is pulmonic egressive, but it can almost as easily be produced ingressively. The exploratory data set contains examples of whistling being used to imitate birds, only in the form of palatal whistling and "digitally assisted" whistling (i.e. finger whistling), possibly because these generate higher oscillation frequencies.

Also, short labial whistling noises can be produced using both glottalic and velaric initiation, again both egressively and ingressively. The exploratory data contain several examples where imitators produce a short whistle with velaric egressive airstream to imitate the impact sound of a drop of water.

Percussive initiation

Percussive initiation does not require an airstream but results instead from an impact between solids, for example when the upper and lower teeth are made to clash or scrape together (Catford 1977: 63).

Percussives occur very rarely in (non-pathological) speech and are not phonologically distinctive in any language. Sands, Maddieson & Ladefoged (1993: 183) observe that, very rarely, an allophonic variant of an alveolar click is a percussive in which "the normal click is quite quiet but the tongue tip makes a forceful contact with the bottom of the mouth after the release of the front click closure". Incidentally, they also mention that this is a "sound sometimes made by speakers of non-click languages trying to imitate the sound made by the shoes of a trotting horse" (ibid.). As we saw in connection with Figure 1, the SkAT-VG exploratory data contain an example of such a "floored", sublaminal percussive, used as part of an impression given by an improvisational actor of "trickling water".

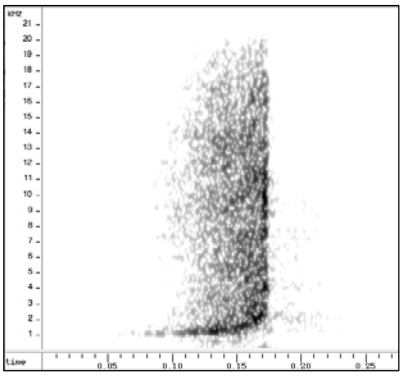

Figure 3. A spectrogram of an actor's impression of the sound of a "whip lash".

The data set also contains an example of a lamino-dental percussive, in which the tongue is shot forward at a

high velocity creating an impact sound as the lamina makes contact with the teeth and the alveolar ridge. This occurred in an improvisational actor's impression of the sound of a "whip lash", shown in the spectrogram in Figure 3. In speech, oral stop sounds are made at the offset of an occlusion by releasing a turbulent airstream through a narrow channel, giving rise to a high energy release burst. By contrast, in the example in Figure 3, the "burst" at 0.17 ms in the spectrogram is created at the onset of the occlusion and is in fact the sound of the impact of the tongue lamina against the teeth.

Conclusion

The observations made during the exploratory phase of the SkAT-VG project have shown that in sound imitations humans can utilize a far wider range of articulations than are used to make phonological distinctions in languages. Also, imitators can utilize sound initiation mechanisms and source types that are not part of the repertoire of their native language(s) and in many cases they utilize mechanisms that are typologically rare (and considered "difficult").

A classification of sound productions is proposed that is based on three basic source types, turbulent, myoelastic and whistled, intersecting with six basic sound initiation mechanisms, pulmonic, glottalic and velaric initiation, both egressive and ingressive. In addition, percussive sounds form a class of their own, being both an initiation mechanism and a source type.

Acknowledgements

I wish to thank Prof. Sten Ternström for helpful advice and comments. This research has been supported by EU-FET grant SkAT-VG 618067.

References

Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

Catford, J. C. (1988). *A practical introduction to phonetics*. Oxford: Clarendon Press.

Eklund, R. (2008). "Pulmonic ingressive phonation: diachronic and synchronic characteristics, distribution and function in animal and human sound production and in human speech." *Journal of the International Phonetic Association 38*: 235-324.

Fant G. (1960). *The Acoustic Theory of Speech Production*. The Hague: Moulton.

Jakobson, R. and Waugh, L. (1979). *The Sound Shape of Language*. Bloomington, Ind: Indiana University Press; and London: Harvester.

Lemaitre, G., Dessein, A., Susini, P. & Aura, K. (2011). "Vocal Imitations and the Identification of Sound Events". *Ecological Psychology, 23*:4, 267-307.

Maddieson, I. & Precoda K. (1990) Updating UPSID. *UCLA Working Papers in Phonetics 74*, pp. 104-111.

Moisik, S. R., Esling, J. H., & Crevier-Buchman, L. (2010). A high-speed laryngoscopic investigation of aryepiglottic trilling. *Journal of the Acoustical Society of America 127 (3)*. 1548-1559.

Pike, Kenneth L. (1943). *Phonetics: a critical analysis of phonetic theory and a technic for the practical description of sounds*. Ann Arbor, MI: University of Michigan Press.

Sands, B., Maddieson, I. and Ladefoged, P. (1993). "The Phonetic Structures of Hadza". *UCLA Working Papers in Phonetics, No. 84: Fieldwork Studies in Targeted Languages*.

Shosted, Ryan. 2006. "Just put your lips together and blow? Whistled fricatives in Southern Bantu." In H. C. Yehia, D. Demolin & R. Laboissiere (Eds.), Proceedings of ISSP 2006, (pp. 565–572). Belo Horizonte: CEFALA.

Stevens, K. N. (1999). *Acoustic Phonetics*. Cambridge: MIT Press.