Gestural undershoot and gestural intrusion – from perception errors to historical sound change

Ioana Chitoran [1]*, José Ignacio Hualde* [2]*, Oana Niculescu* [3]

[1] Université Paris Diderot; CLILLAC-ARP, France
[2] University of Illinois at Urbana-Champaign, USA
[3] University of Bucharest, Romania

ioana.chitoran@univ-paris-diderot.fr, jihualde@illinois.edu, oeniculescu@yahoo.com

Abstract

This paper examines the relationship between sound change and human perception errors that can be attributed either to undershoot of an articulatory gesture or, on the contrary, to the intrusion of an articulatory gesture. We discuss how such phenomena that characterize speech variability can lead to misperceptions that in turn can lead to sound change, resulting in the reorganization of a phonological system (cf. [1], [2]).
In our study we analyze the acoustic manifestation of variable pronunciations in two modern Romance languages (Spanish and Romanian) that are generally assumed to differ in degree of articulatory undershoot. In all modern-day varieties of Spanish, consonant lenition, a result of undershoot, is pervasive. In Romanian it is relatively rare, whereas consonant deletion is widespread in spontaneous speech. In both languages, adjacent vowels tend to coalesce, and unstressed high vowels [i,u] become glides. Gestural intrusion is less common overall. We compare representative examples encountered in spontaneous speech in the two languages, and discuss them in the light of diachronic changes attested.

Index Terms: phonetic variation, target undershoot, weakening, lenition, sound change

1. Introduction – Sound change and phonetic variation

It is understood that sound change happens because variation in speech exists. But for sound change to actually take place, the presence of variation is not sufficient. The auditory system of the listener must crucially intervene to process the signal in such a way that new information is extracted from it and re-encoded in subsequent productions, leading ultimately to a systemic shift. Models of phonetic variation have contributed greatly to explaining sound change (see [3], [4] for a review). John Ohala's phonetic listener-based model ([1], [2]) specifically highlights the listener's role in initiating certain sound changes. According to Ohala, the misperception of the signal by a possibly inexperienced listener is the primary source of sound change. Much of phonetic variation is predictable, due to coarticulation in speech, and as such it is normally factored out by the experienced listener. This type of variation does not typically lead to sound change. Sometimes, however, possibly due to lack of experience or unfavorable listening conditions, the listener fails to correct it and produces the same variation systematically. In this case the listener perceives and subsequently produces a new form, different from the one intended by the speaker. This is the case that Ohala labels *hypocorrection*. The opposite scenario *(hypercorrection)* also exists, and can lead to sound change if the listener mistakenly attributes a certain phonetic event to coarticulation and corrects it. This unnecessary correction results in the production of a new form. Both scenarios involve "misunderstanding in sound change" ([5]). Crucially, both involve a mismatch between production and perception, and that is where certain types of "errors" become directly relevant to the evolution and re-organization of phonological systems over time.

2. Types of variation: listener-perceived vs. speaker-produced

Two major types of speech variation can constitute sources of sound change: the phonetic variation encountered in the signal *produced* by the speaker, and the variation *perceived* by the listener, even if it was not produced. For a review of variation phenomena encountered in speech, see [4]. Here we discuss some examples that are the most relevant for sound change.

One type of variability arises from the interaction between the biological and physical constraints on the speech production and perception mechanisms. A well-known example is the development of tone in East Asian languages ([6]). Variation in speech is not always involuntary, as in this case. It can be determined by the speaker in terms of the intended meaning. Thus, speaking style has important effects on the clarity of speech. More variability is encountered in casual speech as opposed to formal style, or when the information is not new in the discourse as opposed to new information being introduced. Many such examples are found in the data we are analyzing here, from conversational speech in Spanish and Romanian. Finally, another type of variation, relatively less studied so far, is due to non-linguistic factors, related to the identity of the speaker (e.g., age group, social class), their emotional state and attitude.

Having considered the relevant types of phonetic variation that are typically related to sound change, we now turn to a description of the actual phonetic phenomena involved.

3. Articulatory variation

In this section we focus on the articulatory details of two types of phonetic phenomena conducive to sound change: gestural target undershoot and gestural intrusion.

An articulatory target in speech is said to be "undershot" when that target is not fully achieved, but rather approximated. This means that an active articulator (e.g., the tongue, the lips) begins its movement toward the target location (e.g., the teeth or the palate), but does not achieve the degree of constriction specified for it. The phenomenon is commonly encountered in the evolution from Latin to some modern Romance languages.

(1) Lenition in the evolution from Latin to Romance

	/p/	/t/	/k/
Latin	sa[p]ēre	vī[t]a	amī[k]a
Italian	sa[p]ere	vi[t]a	ami[k]a
Spanish	sa[β]er	vi[ð]a	ami[ɣ]a
French	sa[v]oir	vie	amie

Stop lenition is absent in Italian, occurs in Spanish, and in French it can advance further, to full deletion.

Gestural intrusion is, articulatorily, the exact opposite of undershoot: an articulator moves beyond its specified target. A simple example involves the labio-velar glide [w]. Both the labial and the velar constrictions are approximated in the production of the glide. Occasionally children will produce a full velar constriction instead of an approximated one, thus producing a velar stop, as in *Ioana* i**[w]**ana > i**[gw]**ana. In sound change, this "glide strengthening" phenomenon is well attested in the evolution of Germanic words to Modern French and Spanish: **[w]**adja > Gallo-Roman **[gw]**adja (> Modern French **[g]**age) 'wage' ([7]); ***[w]**erra > Sp. **[gwe]**rra, 'war.

4. Data

4.1. Spanish

We have examined a sample of conversational Spanish speech from the corpus *Glissando* ([8]). *Glissando* is a corpus of Spanish and Catalan that includes studio-quality recordings of speech from both professional speakers, such as radio announcers, and regular speakers. All speakers were recorded engaged in task directed conversations and informal dialogs. In addition, the professional speakers were recorded reading the news. This corpus thus provides excellent data for the study of style-conditioned variation in pronunciation. The Spanish portion of the corpus was produced in Valladolid with speakers of Northern Castilian Spanish. For this paper, we have selected three short samples from the informal dialog subcorpus.

Spanish pronunciation is characterized by pervasive consonant reduction, especially in intervocalic and syllable-final position. The voiced stops /b d g/ are systematically realized as approximants in all styles in intervocalic position, and in some postconsonantal contexts. Approximant realizations of /b d g/ vary substantially in degree of constriction, including fully vocalized productions. The most reduced pronunciations may trigger the recategorization of lexical items as not including any consonantal constriction. Thus it is recognized that participles in /-ado/ have a common informal variant in /ao/, e.g. *cantado* [kanˈtaðo] or [kanˈtaọ] 'sung'.

The voiced obstruents /p t k/ may also be realized as voiced approximants in intervocalic position, although this phenomenon lacks the systematicity shown by the allophony of /b d g/. This weakening phenomenon is much more common in informal styles and shows considerable variation among speakers ([9]).

Phonologically, Spanish lacks voiced fricatives, but /s/ and to some extent also /f θ/ may also be realized as voiced in intervocalic position.

There is considerable variation in the realization of coda consonants. In the variety on which we focus here, Northern Castilian Spanish, the coda obstruents /k d/ (other obstruents are uncommon in coda position) may be realized as voiceless fricatives or may be deleted. Syllable-final /s/ is frequently aspirated and lost in many Spanish dialects, both in southern Spain and in Latin America, but this weakening tendency is less strong in the Northern Castilian dialect.

Regarding vowels, Spanish differs from its close relatives, Portuguese and Catalan, in having only relative small differences in quality between stressed and unstressed vowels ([10]). Sequences of vowels inside words or a across word boundaries are subject to coalescence (e.g. [11]).

To illustrate these phenomena, we consider a few representative examples extracted from the *Glissando* informal conversation subcorpus. The first two examples are from the same male speaker (m20s).

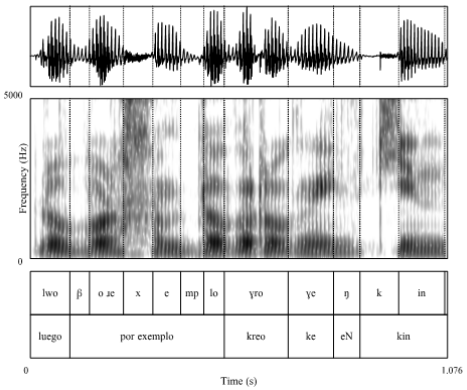

Figure 1. *Luego por ejemplo creo que en quin(to año...*)
/ˈluego por eˈxemplo ˈkreo ke en ˈkin(to ˈaɲo)/
[lwoβoɹexemploɣroɣeŋkin(toaɲo)]
'then, for instance, I believe that in the fifth year'

In figure (1) the word *luego* /luego/ 'then' is produced as the reduced form [lwo]. This is a very high frequency word and arguably the monosyllabic realization may represent a distinct lexical form for this item. Another notable reduction in this fragment is the realization of /p/ in *por* and both instances of /k/ following a vowel (across a word boundary) in *creo que* /kreo ke/ as voiced approximants. The vowel sequence in /kreo/ is also reduced to a single vowel.

In figure (2) there is no spectrographic evidence for the voiced obstruents in *tú dices* 'you are saying' and *sobre todo* 'especially'. It is at least possible that we are dealing here too with lexicalized reduced variants lacking a voiced obstruent. Notice also the voicing of intervocalic /z/, even when it represents an underlying sequence /s#s/ across word boundary.

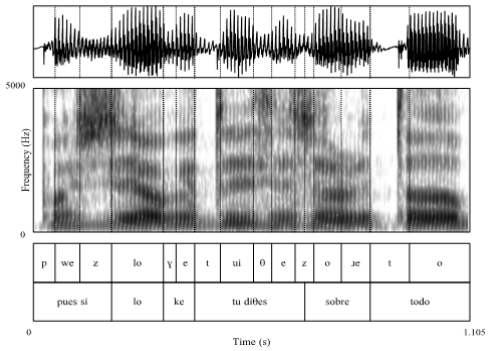

Figure 2: *Pues sí, lo que tú dices, sobre todo para...*
/pues ˈsi lo ke ˈtu ˈdiθes sobre ˈtodo paɾa/
[pwez:ˈloɣeʈuiθ̟ez:oɹeto:paɹa:]
'ok yes, what you are saying, especially for...'

The example in (3) (by speaker f22s) illustrates variation in the realization of intervocalic stops.

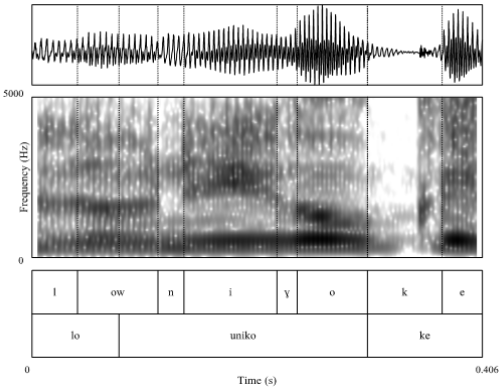

Figure 3: lo único que (f22s)
/lo ˈuniko ke/ [lowniɣok̟e]
'the only thing is…'

The first instance of /k/ in *único* is realized as a voiced approximant. The second /k/, on the other hand, is only partially voiced. Most of the occlusion presents phonetic voicing, but voicing is extinguished before the release, which is voiceless. As mentioned, there is substantial variation among speakers in the weakening of intervocalic voiceless stops.

4.2. Romanian

Data were recorded from 8 speakers (4 male, 4 female). The recordings were done in a quiet room, directly onto a laptop computer, with a stand-mounted microphone (Behringer B1), attached to an external audio interface (M-Audio Fast Track). Participants were required to talk about their previous summer. They each received a form consisting of 6 circularly displayed questions, so that no order is imposed when performing the task. They had to fill in the forms with key words, not full sentences (for more details, see [12]). 15 min were given prior to the elicitation of (semi)spontaneous speech. The interview always started with the same question(s) "What do you think of the structure of this interview? Where would you like to start?" The recordings were transcribed and annotated in PRAAT [13]. In this paper we analyse data from three speakers, two female and one male, age range 26-32. The average time of the recordings is 45 min. Around 3.5 min were analysed for each speaker.

Contrary to the Spanish data, consonant weakening is practically absent in Romanian. The most common phenomena encountered involve vowels, and occur predominantly across word boundaries. They include glide formation, vowel coalescence, vowel blending, vowel deletion, especially of unstressed vowels, full syllable deletion (VC or CV). When consonants are affected, they are acoustically absent. For example they can be deleted acoustically in word-final position before a consonant-initial word. Consonant clusters can be reduced. Only one speaker (mI) has a few instances of what could be considered weakening or lenition, which are widespread in the Spanish data. Figures 4-6 are all from this speaker.

Full syllable deletion is observed in figure 4, where three CV syllables are absent in a short portion at the beginning of the utterance.

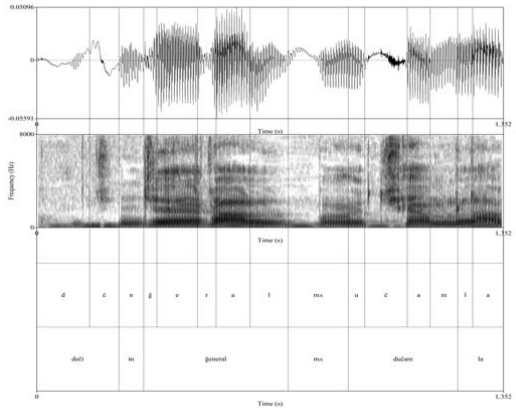

Figure 4: d(e)ci în ge(ne)ral mă (du)ceam (mI)
/ˈdeč̟ʲ ɨn ǧeneˈral mʌ duˈčam/
[dčnǧeralmʌčam]
'so, generally, I went…'

The vowel in the very first syllable, [de], is deleted, although it is stressed, but the word *deci* is a discourse marker signaling continuity with the previous utterance, and does not carry new information.

A drastic example of consonant deletion is observed in figure 5, where *un moment* 'one moment' is reduced to [oen]. The oral constriction of all the nasal stops is absent. The final [t] of *evident* is not realized acoustically before the following stop. Nasal place assimilation occurs in *m-am gândit* 'I thought' (ma[ŋ]â..). The glide of the personal pronoun *io* is also deleted.

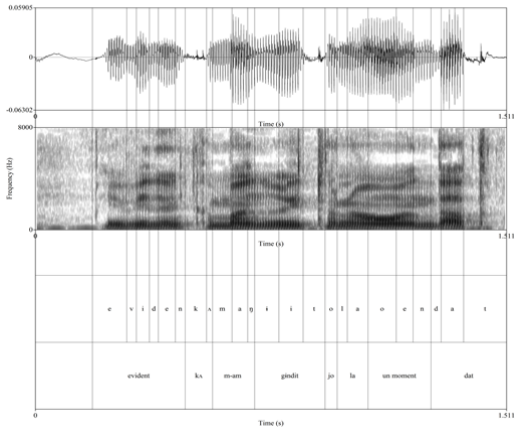

Figure 5: evident că m-am gândit io la un moment dat
/evi'dent kʌ mam gɨn'dit jo la un mo'ment dat/
[evidenkʌmaŋɨitolaoendat]
'Evidently, I thought at one point…'

Finally, a rare example of what could be considered consonant weakening is shown in figure 6. The noun *lucruri* 'things' is realized as [luhu]. The medial stop-liquid cluster is reduced to a glottal fricative.

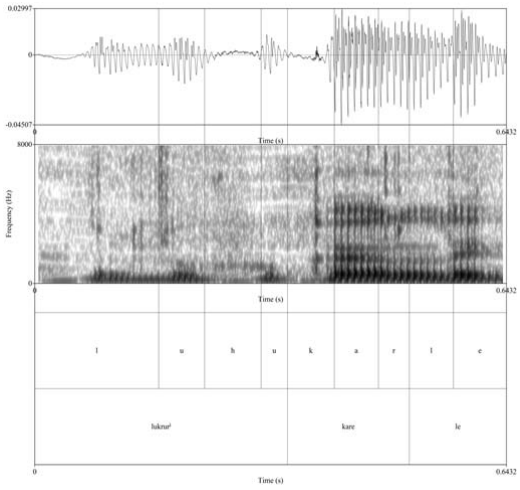

Figure 6: lucruri care le
/'lukrurʲ 'kare le/ [luhukarele]
'… things that…'

Examples involving vowels include elision, as in the phrase *vreau să zic că* 'I mean to say that…', realized as [vreszikʌ] (speaker fD). Vowels across word boundaries often coalesce and are realized as glides. The following types of examples are extremely common. Speaker fR: *cu aşteptările* 'with the expectations' realized as [k**wa**ştept**ʌ**rile]; *nişte articole* 'some articles' realized as [niš**tea**rtikole].

5. Discussion

The comparison of continuous speech data from Spanish and Romanian reveals similar phenomena affecting vowels, and considerable differences with respect to consonants. In both languages, adjacent vowels inside and across word boundaries are prone to coalescence, high and mid vowels tend to be realized as glides. Stressed and unstressed vowels do not differ much in quality. With respect to consonants, however, a major difference is observed. Spanish stops are very sensitive to weakening, which may range from spontaneous voicing to incomplete constrictions (target undershoot). The most common tendencies are: voiceless stops tend to be realized with voicing throughout, and voiced stops tend to be realized as approximants. In our Romanian data, spontaneous voicing is not observed, and undershoot only marginally. Instead, consonants are often unrealized acoustically, which suggests that, rather than having incomplete constrictions, they lack an articulatory release strong enough to produce an acoustically visible (and audible) release burst. Interestingly, then, style-conditioned variation affects primarily the occlusion phase in Spanish, while in Romanian it affects mostly the consonantal release. Such differences in phonetic effects are attributed to cross-linguistic differences in coarticulation ([14]).

When we compare synchronic variation in conversational speech with sound change in the same language, another interesting difference becomes relevant between Spanish and Romanian. Consonant weakening, a major sound change in Spanish, has persisted, and is attested in synchronic style-conditioned variation. The *Glissando* corpus contains numerous examples of weakening of intervocalic consonants. Historically, stops underwent lenition in the same context, as well as between approximants: Lat. ca[p]ra > Sp. ca[b]ra > ca[β]ra. In Romanian none of the synchronic changes correspond to attested historical ones. However, one detail that stands out is the vulnerability of consonant releases in conversational speech. Stop releases are the primary information to place of articulation. If stops tend to be unreleased or weakly released before another consonantal constriction, an expected consequence is perceptual confusion of stop place of articulation. Historically this corresponds to sound changes affecting consonant place of articulation in the relevant environments. Such changes are indeed attested in Romanian, precisely in stop-stop sequences: [kt] becomes [pt] in Lat. *octo*, *lacte* > Rom. *opt*, *lapte* 'eight', 'milk', [gn] becomes [mn] in Lat. *lignu*, *pugnu* > Rom. *lemn*, *semn*, 'wood', 'sign'. Contrary to Romanian, in Spanish such consonant clusters have merged into a palatal consonant, as in *noche* and *leño*, respectively. The Romanian scenario is more likely perceptually-based, while the Spanish one is based on coarticulation in production.

6. Conclusions

The comparison of conversational data in two Romance varieties has revealed interesting differences in consonantal variation, and similarities in variation involving vowels. In Spanish, the most pervasive type of variation, weakening, closely reflects the same diachronic change. In Romanian, the relationship between synchronic variation and sound change is more subtle, but the comparison suggests so far a particular sensitivity of the consonantal release phase. To complete this analysis, perception studies are critical, testing the perception of relevant portions of data from conversational speech. Testing of human and machine perception is equally informative for these studies.

7. Acknowledgements

O. Niculescu acknowledges SOP HRD, 4 Uniunea Europeană Guvernul României, Fondul Social European POSDRU 2007-2013, contract no. SOP HRD/159/1.5/S/136077.

8. References

[1] J.J. Ohala. "The listener as a source of sound change", In C.S. Masek, R.A. Hendrick, and M.F. Miller (Eds.) *Papers from the Parasession on Language and Behavior*. Chicago: Chicago Linguistic Society, pp. 178-203, 1981.

[2] J.J. Ohala. "The phonetics of sound change", In C. Jones (Ed.), *Historical Linguistics: Problems and Perspectives*. London: Longman, pp. 237-278, 1993.

[3] I. Chitoran. "The nature of historical change". In A.C. Cohn, C. Fougeron, M. Huffman (Eds.) *The Oxford Handbook of Laboratory Phonology,* pp. 311-321. Oxford University Press. 2012.

[4] J. Harrington. "The relationship between synchronic variation and diachronic change". In A.C. Cohn, C. Fougeron, M. Huffman (Eds.) *The Oxford Handbook of Laboratory Phonology,* pp. 321-332. Oxford University Press. 2012.

[5] W. Labov. *Principles of linguistic change, vol. 1: Internal Factors*. Oxford: Blackwell. 1994.

[6] J-M. Hombert. 1979.

[7] P. Delattre. "Stages of Old French phonetic changes observed in Modern Spanish". *Publications of the MLA* 61(1), pp. 7-41, 1946.

[8] J.M. Garrido et al. "Glissando: a corpus for multidisciplinary prosodic studies in Spanish and Catalan". *Language Resources and Evaluation* 47, pp. 945-971. 2013

[9] J.I. Hualde, M. Simonet, and M. Nadeu. "Consonant lenition and phonological recategorization". *Laboratory Phonology* 2, pp. 301-329. 2011

[10] M. Nadeu. "Stress- and speech rate-induced vowel quality variation in Catalan and Spanish". *Journal of Phonetics* 46, pp. 1-22. 2014.

[11] J.I. Hualde, M. Simonet, and F. Torreira. "Postlexical contraction of non-high vowels in Spanish". *Lingua* 118, pp. 1906-1925. 2008

[12] O. Niculescu. Hiatul în limba română: O abordare optimalistă. University of Bucharest PhD dissertation. Forthcoming.

[13] P. Boersma and D. Weenink. PRAAT: doing phonetics by computer. http://www.praat.org/. 2015.

[14] S. Manuel. "Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts". In W. Hardcastle and N. Hewlett (Eds.) *Coarticulation: Theory, data, and techniques*, pp. 179-198. 1999.