DOI: 10.17469/O2101AISV000002

DANIEL RECASENS

Phonetic factors contributing to the inception and evolution of sound change

This paper uses experimental evidence for showing that, depending on the case, sound changes may be triggered primarily by either articulatory variation (as for changes occurring through segmental weakening or strengthening) or by acoustic equivalence (as for vowel nasalization or for segmental substitutions involving syllable-final stops of different places of articulation). It also argues for a multiple evolutionary pathway in the case of specific sound changes such as dark /l/ elision or the palatalization of Latin /kt, ks/ in Romance.The role of word prominence in vowel assimilations and dissimilations and how sound changes originate in the individual are also looked into.

In recent times more attention has been paid to how sound changes spread through the lexicon and the linguistic community (Labov, 1994; Phillips, 2006; Wang, 1969) than to the phonetic factors which contribute to the origin of sound change. Fortunately the situation is improving thanks to the pioneering work of Ohala (see section 1.2) and an increasing number of books and articles on this research topic (Solé, Recasens, 2012; Yu, 2013). This scenario is in clear contrast with the one existing during the last decades of the XIX century and the first decades of the XX century, as revealed by outstanding publications on the articulatory causes of sound change which appeared at that time (Grammont, 1933; Millardet, 1910; Rousselot, 1897-1901). One reason for the present lack of interest in the phonetic causes of sound change is to be sought in the special emphasis that structuralist and generative linguists have put into the study of synchronic phonology much to the exclusion of diachronic aspects and phonetic variation. Another reason may lie in the difficulty involved in identifying those articulatory and acoustic characteristics which render the replacement of a linguistic sound by another feasible and in determining the degree to which speakers become more or less sensitive to spectral changes which may occur during the phonetic realization of phonemes. Keeping these aspects in mind, the present study deals with the sound change inception problem by looking into the following issues: whether sound changes have an articulatory and/or an acoustic origin (section 1) and their implementation follows a single or a multiple pathway (section 2); the contribution of articulatory/acoustic prominence to sound change (section 3); how sound change originates in the individual (section 4).

1. *Articulation- vs acoustic-based sound changes*

Linguists and philologists working on sound change during the first half of the XX century thought that sound shifts were essentially articulation-based. The general principle underlying this belief was that listeners may fail to identify a given phoneme intended by the speaker whenever subject to prominent articulatory and spectral changes in specific contextual, positional and/or prosodic conditions. According to an alternative acoustico-perceptual based approach, on the other hand, a number of sound changes are triggered by acoustic equivalence rather by articulatory variation (Ohala, 1981). The rationale behind this assumption is that two phonemic units are prone to be confused perceptually whenever exhibiting similar spectral cues in favorable segmental contexts and positions within the word and the utterance. A critical evaluation of these two explanatory hypotheses of sound change is presented next.

1.1 Articulatory origin

Most sound changes originate whenever listeners fail to compensate for coarticulation. Context-dependent coarticulatory effects may give rise to articulatory changes in the phonetic implementation of a given phoneme which may lead listeners to believe that a new phoneme has been produced by the speaker. An illustrative example is the raising of stressed /a/ to a mid front vowel before an (alveolo)palatal consonant in Early Romance, as in the developments Latin [ˈakse] AXE "axis" > *[ˈajse] > Catalan [e(j)ʃ] and Latin [ˈfakto] FACTU "done" > *[ˈfajto] > Catalan [fet], where [j] in the intermediate phonetic variants *[ˈajse] and *[ˈfajto] corresponds to a lenited realization of the original syllable-final velar stop. In this particular case productions of /a/ involving some tongue body fronting and raising, and consequently an increase in F2 frequency, induced by following [j] must have been heard as realizations of the mid front vowel phoneme by the listeners. Another clear example of an articulation-based sound change is the palatalization of a dentoalveolar consonant by an adjacent front vowel or glide, as in the above development Latin [ˈakse] AXE "axis" > *[ˈajse] > Catalan [e(j)ʃ] where the glide has caused following /s/ to turn to a palatoalveolar fricative after which it has stayed or dropped (both variants [ejʃ] and [eʃ] occur in Catalan depending on dialect).

In partial disagreement with linguists and philologists from the XIX and XX centuries, data reported in sections 1.1.1 through 1.1.3 show that there are sound changes which appear to be acoustically driven and thus are not triggered by articulatory variation.

1.1.1 Syllable-final stops

The exchange between oral stops of different places of articulation in syllable-final position whether word-internally or word-finally is prone to be caused by acoustic equivalence (see also Gauchat, 1925). Thus, for example, the spectral similarity between the stop bursts and the vowel formant transitions for /k/ and /t/ next to a front vowel may account for the changes /t/ > [k] and /k/ > [t] in this vowel context whenever the originary stop occurs at the end of the syllable and is thus not too perceptually salient. The former change has taken place in Standard Catalan *ànec* ANATE "duck" and *préstec* PRAESTITU "loan" where /t/ in the Latin forms was replaced by [k], and the latter in the dialectal Catalan variants *coït* CULICINU "mosquito" and *présset* PERSICU "peach" where /k/ in the Latin lexical items shifted to [t] (the lexical cognates for *coït* and *présset* in Standard Catalan are *coïc* and *préssec* and therefore have a velar stop, as in Latin). The fact that /t/ and /k/ may be exchanged in the two directions, i.e., from dental to velar and from velar to dental, suggests that the replacements in question are triggered by spectral equivalence rather than by variations in closure fronting. Indeed, an articulatory account of this reversable change is feasible for /k/ > [t] (the velar may be confused with the dental when fronted to a [c]-like articulation, e.g., before a front vocalic segment) but not for /t/ > [k] (no tongue closure retraction is expected to occur in this case, e.g., before a back rounded vocalic segment).

1.1.2 Vowel nasalization

Changes in vowel quality induced by nasalization such as Old French /ỹ/ > [œ̃] (*un* UNU "one") and /ĩ/ > [ɛ̃] (*vin* VINU "wine") have also been triggered by spectral modifications. While Grammont (1933) suggested that those changes were articulation-based, Beddor (1983) has argued convincingly that they arise when nasal formants are added to the first vowel formant region thus causing higher oral vowels to be heard as if they were lower (and low vowels to be perceived as if they were higher) when nasalized. Recent evidence indicates that the spectral changes associated with nasalization may be enhanced or attenuated by changes in articulatory configuration which may occur during the production of nasalized vowels (Carignan, Shosted, Shih & Rong, 2013).

1.1.3 Sonorants

A third example of an acoustically driven sound change is the shift from clear /l/ to an alveolar tap ([ɾ]), as in the case of dialectal Spanish *alto* becoming *arto* ALTU "high". It may be argued that this sound change is conditioned by segmental shortening (the alveolar tap is shorter than the clear alveolar lateral) and gestural reduction (whenever too short, the tongue front may fail to make a complete closure during the production of clear /l/). Whether these articulatory maneuvers apply or not, the replacement of /l/ by an alveolar tap appears to be finally determined by the fact that the two consonants share a relatively high F2 about 1600 Hz or higher and are thus spectrally similar.

1.2 Acoustic origin

While, as shown in section 1.1, several sound changes which were traditionally thought to be articulation-based appear to have an acoustic origin, there also are sound changes which have been attributed to acoustic equivalence and are more prone to have an articulatory origin instead. As exemplified next, evidence for the articulatory origin of specific sound changes is to be sought in the existence of intermediate variants between the input and output forms which may have been available at different historical periods.

1.2.1 Syllable-initial stops

Segmental substitutions involving syllable-initial stops of different places of articulation may occur in very specific vocalic contexts, as for example /k/ > [p] before a back rounded glide (Indo-European **ekwos* > Classical Greek *hippos* "horse"; Latin [ˈkwattuor] QUATTUOR > Romanian *patru* "four", Sardinian *báttoro*) and palatalized /p/ > [(t)ʃ] (Latin *sap*[j]*a* SAPIAM > Fr. *sache* "know, 1st person, subjunctive tense"; Latin *p*[j]*eno* PLENU > Portuguese *cheo* "full"). These sound changes have been attributed to the existence of similar burst spectra and formant transitions for the originary and outcoming consonants (Ohala, 1978; 1989). Indeed, both the burst and the F2 vowel transition locus are low frequency for labial and back velar stop consonants, and high frequency for the palatalized labial stop and the palatoalveolar affricate (also for /t/).

A closer analysis of these phonetic developments reveals, however, that, at least in a subset of languages (Hickley, 1984), they may have involved an increase in glide constriction narrowing and thus an articulatory mechanism. Under this assumption the phonetic derivation for palatalized /p/ would then be /pj/ > [pç] > [pc] > [ptʃ] > [(t)ʃ], which is consistent with lexical variants exhibiting the intermediate stages [pç], [pc] and [ptʃ] (Engadin Rhaetoromance [ˈsapca] SAPIAM, Occitan [ˈsaptʃa] SAPIAM, Romanian [ˈpçatɾə], [ˈ(p)catɾə] PETRA "stone"; Lausberg, 1970: 398; Rankin, 1974). Given this scenario, it may be suggested that the replacement of /kw/ by [p] has been triggered by an increase in labiovelar constriction degree (/w/ > [ɸ]) and therefore has proceeded through the pathway /kw/ > [kɸ] > [kp] > [p].

1.2.2 /l/ vocalization

The vocalization of /l/ may yield the outcomes [w] or [j] depending on whether the lateral is dark or clear, respectively. Thus, the replacement of dark /l/ by [w] accounts for the dialectal Catalan variants *a*[w]*ba* ALBA "dawn" and *esca*[w]*far* EXCALEFACERE "to heat", and the substitution of /l/ by [j] for Italian *amp*[j]*o* derived from Latin AMPLU. According to the acoustic equivalence hypothesis (Ohala, 1974), these changes are associated with whether the /l/ steady-state period has a low F2 frequency about 800-1000 Hz ([ɫ] > [w]) or else a high F2 about 1600-1800 Hz ([l] > [j]).

There is however evidence for an articulation-based interpretation of these two consonant vocalization processes in Romance. Thus, it has been hypothesized that the replacement of clear /l/ by [j] in the tautosyllabic clusters /pl/ and /kl/ has not taken place directly but through the intermediate realizations [pʎ] and [kʎ] which are still available in several Romance dialects, e.g., Ribagorçan Catalan [kʎaw] CLAVE "key", [ˈpʎɔwɾe] PLOVERE "to rain", Francoprovençal areas [ˈpʎema] PLUMA "feather", [kʎa] CLAVE (Repetti, Tuttle, 1987). In particular, the development /kl/> [kʎ]> [kj] involves two consecutive production mechanisms: blending between the tongue front gesture for the alveolar lateral and the dorsal gesture for the velar stop into the alveolopalatal lateral realization [ʎ], followed by delateralization of [ʎ] into [j] through central tongue contact loss. Regarding the change [ɫ] > [w], there is articulatory evidence indicating that dark /l/ vocalization has been induced by apical contact loss, which contributes to lower the F2 frequency down to 600-800 Hz thus rendering the alveolar lateral spectrally similar to /w/ (Lin, Beddor & Coetzee, 2014). This articulatory account is also consistent with the fact that [ɫ] vocalization is prone to occur before labial and velar consonants (Gevaudanès Occitan [ʻawbo] ALBA, [faw'ku] FALCONE), where the lateral may exhibit less apicoalveolar contact and a lower F2 than before consonants of other places of articulation (Hardcastle, Barry, 1985; Scobbie, Wrench, 2003).

In order to seek for a more thorough understanding about whether [ɫ] vocalization is acoustic- or articulation-based we conducted a perception experiment where several Catalan subjects were asked to identify as /l/ or as /w/ several tokens of [ɫ] embedded in [VɫCV] stimuli excised from real speech sequences produced by several speakers of Majorcan Catalan where the alveolar lateral has a strongly dark realization (Recasens, Espinosa, 2010a). A relevant aspect about the perception stimuli was that, since they had been produced with an artificial palate in place, we could test the effect of changes in F2 frequency and in alveolar contact size in the phonemic categorization of [ɫ]. As exemplified in Table 1, the perception stimuli subject to identification had the following spectral and articulatory properties: their F2 could range between 650-750 Hz (more /w/-like) and 800-950 Hz (less /w/-like); they could be articulated with minimal or maximal apicoalveolar contact if showing, respectively, 0 to 3 or 6 to 8 'on' electrodes at any of the alveolar rows of electrodes on the artificial palate (the maximal number of electrodes on a given alveolar row is 8). Given these F2 frequency and alveolar contact ranges, the [VɫCV] stimuli put to test were characterized by either a low or a high F2 and minimal or maximal alveolar contact. The consonant following [ɫ] could be the stops /b, d, k/ and the alveolar fricative /s/ (only /ld/ stimuli specified for the maximal alveolar contact condition were included since /l/ is always articulated with a full closure before /d/ in Catalan). The initial testing hypothesis was that, if spectrally dependent, [ɫ] would be heard as /w/ not only in the case of stimuli showing minimal alveolar contact and a low F2 but also a low F2 and maximal alveolar contact; on the other hand, if vocalization was mainly dependent on alveolar contact loss and thus articulation based, [ɫ] would be heard as /w/ when produced with minimal alveolar contact independently of whether F2 was low or high.

The identification test yielded highest /w/ identification percentages for the low F2 – minimal alveolar contact stimuli, more so when /l/ was followed by /b/ (57%) than by /d, s, k/ (15%-35%). The /w/ identification percentages for all clusters in the three other acoustico-articulatory conditions were generally below 20%. While these identification results do not seem to provide clear support for either of the two explanatory hypotheses, there are reasons to believe that the articulatory account has higher predictive power than the acoustico-perceptual one: alveolar contact loss appears to cause F2 to lower to a greater extent than other articulatory maneuvers such as tongue body retraction thus giving rise to a /w/-like spectral configuration more effectively and this frequency lowering effect is especially prone to occur before labial consonants which is where dark /l/ vocalization occurs most often.

Table 1 - *Articulatory and spectral conditions for the identification test of dark /l/ as /w/ in the cluster/lb/*

speaker	linguoalveolar contact		F2 frequency	
1	low	0 ‘on’ electrodes	low	657 Hz
2	low	0 ‘on’ electrodes	low	666 Hz
3	low	1 ‘on’ electrodes	high	882 Hz
1	low	0 ‘on’ electrodes	high	824 Hz
4	high	8 ‘on’ electrodes	low	755 Hz
4	high	8 ‘on’ electrodes	low	762 Hz
3	high	8 ‘on’ electrodes	high	898 Hz
4	high	7 ‘on’ electrodes	high	827 Hz

1.2.3 Velar softening

Another sound change process which may be accounted for on an articulatory basis despite having been attributed to acoustic equivalence is velar softening, i.e., the shift from a front allophone of /k/ before a front vowel or glide to the palatoalveolar affricate [tʃ]. By virtue of this sound change Old English [kin] became [tʃɪn] in Modern English and Latin [ˈkɛnto] CENTU “one hundred” turned to [ˈtʃɛnto] in Italian. According to the acoustico-perceptual hypothesis, this change is rendered possible due to the similarity between the 2500-3500 Hz spectral peak for the stop burst of front /k/, which is typically articulated at the palatovelar zone, and the frication noise of the palatoalveolar affricate /tʃ/. Spectral similarity ensures that front /k/ will be heard as /tʃ/ whenever the stop burst is sufficiently long and/or intense as the central lingual constriction at stop release undergoes considerable narrowing in anticipation of the front vowel or glide (Guion, 1998). According to the articulation-based account, on the other hand, velar softening may take place if closure location for the front velar stop shifts from the postpalato-velar zone to the palatal and even alveolopalatal zone thus giving rise to [c] (Rousselot, 1897-1901). This (alveolo)palatal stop realization, which may be articulated at a very close location to /tʃ/, may then be identified as the palatoalveolar affricate whenever its burst is sufficiently prominent. There are several events in support of an articulation-based account of velar softening through the intermediate realization [c] in Romance: a considerable number of Romance languages (Romanian and Majorcan Catalan, as well as Francoprovençal, Northwestern French, Northern Italian and Rhaetoromance dialects) have the allophone [c] of /k/ (also the allophone [ɟ] of /g/) in their consonant system; velar softening may occur not only before a front vocalic segment but also before /a/ and word finally where it can only be possibly triggered by an (alveolo)palatal realization of the velar stop which may indeed be found in those two vowel and word conditions in some of the Romance dialects listed above (Recasens, 2011).

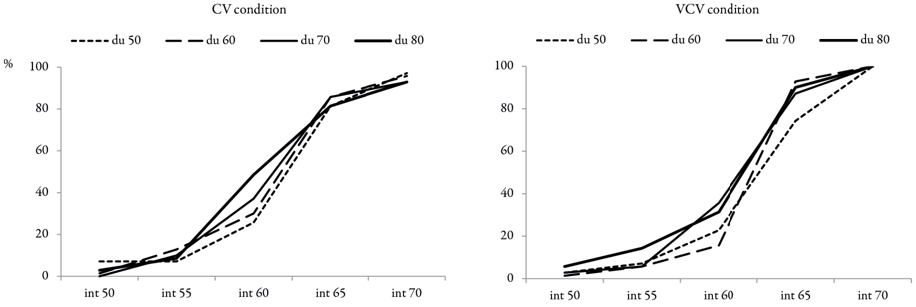
Two perception experiments (see (a) and (b) below) were carried out in order to investigate whether [c] may be heard as /tʃ/ and, if so, which acoustic cues are primarily responsible for velar softening.

(a) Real speech [cV] excerpts excised from a subset of [V#cV], [Vc#V], [VcV] and [#cV] sequences with the vowels /i/ and /a/ produced by several Majorcan Catalan speakers were presented for identification as /k/ or as /tʃ/ (in Majorcan Catalan, [c] is an allophone of /k/ occurring before /i, e, ɛ, a, j/ and word finally). According to the perceptual identification results, [c] is prone to be heard as /tʃ/ before /i/ rather than before /a/ (affricate responses amounted to 40-70% in the former vowel condition and to less than 40% in the latter), which is in line with velar softening occurring mostly before front vocalic segments in the world's languages (Recasens, 2014).

(b) Since the frequency characteristics of the stop burst and the vowel formant transitions for [c] approach those for /tʃ/ independently of segmental context and word position, it was hypothesized that the identification of [c] as /tʃ/ before front vowels and [j] ought to depend on an increase in burst duration and/or intensity associated with the articulatory and aerodynamic factors referred to above. In order to elicit which one of these two acoustic characteristics would yield higher affricate identification percentages, Catalan listeners were asked to identify as /k/ or /tʃ/ a subset of the [#ci] and [i#ci] stimuli used in the previous perception test whose burst duration and intensity values were more appropriate for the affricate and thus ranged between 50 ms and 80 ms and 50 dB and 70 dB, respectively. The identification results show that the intensity level is the most prominent acoustic cue for the identification of [c] as /tʃ/: the affricate is heard about 80%-100% of the time when burst intensity is about 65-70 dB, which is comparable to the frication noise intensity for the affricate. On the other hand, burst duration may also play a role though only when the intensity level is about 60 dB and therefore lies somewhere between the intensity levels for unaspirated [k] and for [tʃ].

The replacement of [c] by /tʃ/ before /a/ and word finally may be achieved through an increase in linguopalatal contact towards the alveolar zone and thus articulatory enhancement (see also Garrett, Johnson, 2010). This tongue contact increase before /a/ (and thus the chances that velar softening applies) is expected to take place mostly when the stop occurs in the phonetically prominent word initial and stressed syllable conditions.

Figure 1 - *Cross-subject affricate identification percentages for [c] as /tʃ/ in the [#ci] and [i#ci] conditions. Percentages are given for five [c] burst intensity levels (int = 50, 55, 60, 65 and 70 dB) and four [c] burst duration values (du = 50, 60, 70 and 80 ms)*



2. *Single vs multiple evolutionary pathways*

It has been traditionally assumed that sound changes proceed along a single pathway. Several changes depend indeed on a single articulatory mechanism which can be identified relatively easily. Thus, articulatory enhancement may account for why word-initial /n, l/ or their geminate cognates /nn, ll/ may be replaced by the alveolo-palatals /ʎ, ɲ/ (Argentinian Spanish [ˈɲuðo] for [ˈnuðo] NODU "knot", Catalan [ʎas] LAQUEU "bow", [aɲ] ANNU "year"). Articulatory reduction, on the other hand, can explain why /ɲ/ may yield [n] in syllable- or word-final position, as shown by the alternation between syllable-initial intervocalic [ɲ] and word-final [n] in the Spanish lexical pair *desdeñar* "treat with disdain" - *desdén* "disdain". A rise in intraoral pressure level conditioned to a greater or lesser extent by articulatory factors explains why oral stop epenthesis may take place in nasal+ fricative clusters (e.g., /ms/ > [mps]) and why syllable-final obstruents are prone to devoice in heterosyllabic consonant clusters (e.g., /zt/ > [z̥t]).

Attention to phonetic detail suggests that several sound changes may proceed through more than one evolutionary pathway depending on the relative prominence of the acoustic characteristics that speaker-listeners use for phonemic identification. This possibility is in line with vowels and consonants being often cued by several cooccurring acoustic characteristics (see Lisker, 1986 for the acoustic properties that cue the stop voicing distinction in the world's languages). This section deals with several phonemic replacements which may be achieved through more than one evolutionary pathway and seeks to determine which acoustic cues are used by listeners for implementing each sound change.

2.1 Dark /l/ vocalization

Languages and dialects may choose between two different pathways for the identification of [ɫ] as /w/, most clearly so when the lateral is preceded by a low vowel:

(a) The misperception of dark /l/ as /w/ whenever the steady-state spectrum for the alveolar lateral exhibits a sufficiently low F2 frequency, i.e., /al/ > [aw]. This is by far the most preferred option in the world's languages (see section 1.2.2).

(b) The integration of the VC transitions as /w/ (/al/ > [awɫ]) followed by [ɫ] elision whenever the apical closure is severely reduced or occurs after voicing offset utterance-finally ([awɫ] > [aw]). This phonetic development occurs in some Romance dialects which still keep the intermediate phonetic variant *aul*, as for [awɫt] ALTU "high" (Romansh) and *aurdeia* 'aldea' "village" (Northern Portuguese dialects).

Spectrographic and perception data reported elsewhere (Recasens, Espinosa, 2010) show that in order for /al/ to be heard as /awl/ the VC transitions must start before the steady-state portion of the alveolar lateral and therefore the tongue dorsum lowering and backing motion for dark /l/ must occur in advance of the tongue tip raising gesture for the consonant. A perception experiment was carried out in order to ascertain whether the identification of dark /l/ as /w/ is likely to proceed either through the pathway /Vl/ > [Vw] when F2 lowering at the consonant steady-state period causes dark /l/ to be perceived as /w/, or else the pathway /Vl/ > [Vwɫ] > [Vw] whenever variations in the relative timing of the VC transitions cause an [w]-like segment to be perceived at /l/ onset. A series of synthetic speech stimuli were built up which differed in the two acoustic parameters as follows: the steady-state F2 frequency values ranged between 650 Hz for a strongly dark realization of [ɫ] and 1050 Hz for a less dark realization of the consonant; the VC transitions ended at four different temporal points: at the onset of the /l/ steady-state period, or at 15 ms, 30 ms and 45 ms before it. Subjects were asked to identify the stimuli in question as /l/ or /w/.

Perceptual identification results show that /l/ is more likely to be identified as /w/ when F2 at the steady-state period of /l/ is lowered from 1050 Hz down to 650 Hz (by 5-45%) than when the offset of the VC transitions is anticipated in time (by 15-35%). Moreover, the latter acoustic characteristic contributes to an increase in the /w/ identification percentages only when the F2 frequency at the steady-state portion of the alveolar lateral is minimal (650-750 Hz). These perception results are consistent with the outcome of the vocalization process being [w] rather than [wl] in the world's languages, and suggest that the vowel transitions may act as a secondary vocalization cue which is used by speakers to enhance the perceptual effectiveness of the steady-state F2 frequency whenever /l/ is strongly dark.

2.2 Dark /l/ elision

Another example illustrating a double evolutionary pathway is the elision of dark /l/ in preconsonantal position mostly after /a/ or a back rounded vowel. It has been assumed traditionally that this change must take place after /l/ vocalization and thus that the final outcome *o*C of /alC/ (Spanish *otro* ALTERU "other", French *haut* ALTU) has been issued through the stages /alC/ > [awC] > [ɔwC] > *o*C where the lateral vocalizes in the first place, /a/ assimilates to following [w] at a later stage, and [w] ceases to be heard due to perceptual similarity with preceding *o* at the end of the derivation. This development is supported by the existence of the intermediate sequence *ou* in lexical forms such as Portuguese *outro* ALTERU.

Another phonetic evolution is also possible, i.e., /alC/ > [ɔɫC] > *o*C, where consonant-dependent anticipatory tongue body lowering and retraction causes lower pharyngeal /a/ to raise to upper pharyngeal [ɔ] after which [ɫ] ceases to be heard due to the spectral similarity with the mid back rounded vowel. Supporting evidence for this evolutionary pathway comes from dialectal variants such as Lombard *molta* MALTHA "malt" (Rohlfs, 1966: 37), as well as the absence of specific intermediate vocalized variants. Thus, for example, Catalan *cop* COLAPHU "blow" must have been issued directly from *colp* since *colp* and *cop*, but not *coup*, appear in old written texts (Coromines, 1980-1988, 2: 888).

2.3 /Vɲ/ > [VjN]

It has been assumed that the decomposition of /ɲ/ into [jN], which occurs mostly syllable-finally, may operate whenever the alveolopalatal nasal exhibits a reduced realization involving a decrease in dorsopalatal contact. At least among a subset of speakers, this change has become phonologized in preconsonantal word final position in Majorcan Catalan. Thus, in this Catalan dialect, /aɲ/ is realized as [ajm] in the sequence *any bo* "good year" and /paɲ/ as [pajn] in the sequence *pany dur* "tough lock", where regressive place assimilation applies to the nasal consonant outcome of the /ɲ/ decomposition process. However, while weakening could certainly account for the reduction of /ɲ/ to [n] (see section 2), it leaves unexplained the presence of the on-glide in phonetic outcomes such as [jm] and [jn] above. An alternative explanatory account for this on-gliding process is that the replacement of /ɲ/ by [jN] has been triggered not by articulatory weakening but by articulatory reinforcement involving an increase in dorsopalatal contact and consequently in the frequency endpoint and the extent of the F2 vowel transitions (/ɲ/ > [jɲ] > [jN]). Several data speak in support of this phonetic development: electropalatographic data for Majorcan Catalan show that syllable-initial /ɲ/ may be realized not only as an alveolopalatal consonant exhibiting a closure location encompassing the alveolar and prepalatal zones but also as a purely palatal consonant which is articulated at the palatal zone exclusively; in other Romance languages, [j] insertion may also take place occasionally before /ɲ/ and other syllable initial (alveolo)palatal consonants without the consonant becoming depalatalized after on-gliding has taken place (see Recasens, 2014 for examples).

In order to elicit whether the change /ɲ/ > [jN] is triggered by articulatory weakening or by articulatory strengthening, the perceptual role of the two following acoustic cues acting in synchrony was submitted to perceptual evaluation in Recasens, Espinosa (2010):

(a) The duration of the VC transitions, which ought to be longer for [jn] and [jɲ] than for [ɲ] and, therefore, could be associated both with a reduced version of [ɲ] (/ɲ/ > [jN]) and with an increase in dorsopalatal contact during the alveolopalatal nasal (/ɲ/ > [jɲ]).

(b) The VC transitions endpoint frequency, which should be higher for [jɲ] (though not necessarily for [jN]) than for [ɲ] and therefore ought to be positively related to an increase in dorsopalatal contact. Indeed, the larger the dorsopalatal contact size, the higher the F2 frequency at the offset of the vowel transitions is expected to be.

In order to find out which one of the two acoustic cues contributes most to /ɲ/ decomposition, a series of LPC-generated synthetic stimuli were built up using a representative production of the sequence /aɲ/ by a native Catalan speaker. The stimuli differed both in the endpoint frequency and duration of the F2 vowel transitions, which ranged between 1800 Hz to 2200 Hz and 70 and 170 ms respectively. Subjects were asked to identify these acoustic stimuli as either /ɲ/ or /jn/. Results show that the /jn/ identification percentages increase with the endpoint frequency (by 10%-50%) rather than with the duration (by 20%- 40%) of the F2 vowel transition; moreover, the duration cue was only perceptually effective when the endpoint frequency amounted to about 2100-2200 Hz and was thus highest. This finding is in support of the hypothesis that an increase in dorsopalatal contact may cause /ɲ/ to be heard as /jn/ since, as pointed out above, this articulatory characteristic and the F2 transitions endpoint frequency are positively correlated. Results also reveal that the duration of the F2 vowel transitions acts as a secondary cue by contributing to an increase in the percentage of /jn/ responses whenever the frequency endpoint exceeds 2000 Hz.

2.4 Palatalization of /kt, ks/

Another relevant sound change which illustrates the multiple vs single pathway is the evolution of /kt/ and /ks/ in the Romance languages. Two derivations (a) and (b) have been proposed traditionally in order to account for the phonetic outcomes /jt/, /tʃ/ and [c] of /kt/ (as in Portuguese *leite* LACTE “milk”, Spanish *leche* and Rhaetoromance [lac]), and /js/ and /ʃ/ of /ks/ (as in French *fraisse* FRAXINU “ash” and Catalan *freixe*):

(a) The velar stop undergoes lenition and vocalization into [j] after which progressive palatalization causes the following dentoalveolar to turn into an (alveolo)palatal consonant (Menéndez Pidal, 1968: 143-144). At a later stage, the outcome [jc] may yield [c] though glide absorption and [c] may give rise to an affricate.

/kt/ > [çt] > [jt] > [jc] > [c] > [tʃ]
/ks/ > [çs] > [js] > [jʃ] > [ʃ].

(b) The two consonants of the cluster /kt/ blend into an alveolopalatal or palatalized stop ([c], [t^j]) and those of the cluster /ks/ into a palatoalveolar or palatalized fricative ([ʃ], [s^j], after which several changes (full palatalization, on-gliding, affrication) may take place. Two slightly different formulations of this development have been proposed: Hall's in (a') (Hall, 1974) and Thomsen's in (b') (Thomsen, 1876).

(a') /kt/ > [xt] > [çt] > [tʲ] > [tʃ, jt]
/ks/ > [xs] > [çs] > [sʲ] > [ʃ, js]

(b') /kt/ > [tʲ:] > [jt, c, tʃ]
/ks/ > [sʲ:] > [js, ʃ].

At least regarding originary /kt/, we would like to propose that the two phonetic derivations, not just one, may have operated in Romance. Indeed, the change /kt/ > [çt] > [jt] looks more plausible than /kt/ > [c] > [jt] in the case of dialects and languages where the final outcome is [(j)t] such as Portuguese, Old Spanish and Catalan, the main reason being that the alveolopalatal stop [c] is not prone to develop an on-glide due to its fast and relatively short vowel transitions. On the other hand, the derivation involving gestural blending is preferable in the case of dialects and languages where the final outcome is [c] (Rhaetoromance, Northern Italian dialects). It should also be noticed that, as the developments above show, the final outcome [tʃ] may derive from [c] in the two derivations and thus whether the (alveolo)palatal stop is generated from [jt] through progressive palatalization (as in Spanish) or through blending (as in Rhaetoromance).

As for /ks/, manner requirements for /s/ render unlikely the implementation of a blending process between the stop and the following fricative consonant the reason being that the tongue dorsum needs to be lowered somewhat for the passage of airflow through a central slit. Therefore, only the vocalization pathway referred to in (a) above and the following developments appear to be possible in this case: /ks/ > [çs] > [js] > [jsʲ] > [js] (French, Occitan) and /ks/ > [çs] > [js] > [jʃ] > [ʃ] (Catalan, Portuguese, Old Spanish).

3. *The contribution of articulatory/acoustic prominence*

This section explores briefly the relevance of positional factors in sound change implementation. Articulatory and acoustic data reveal that consonants are produced with more tongue contact and vowels become more open word-initially than word-internally, and that phonetic segments occurring in stressed syllables are more prominent than those embedded in unstressed ones (Fougeron, Keating, 1997).

In a recent book (Recasens, 2014) we analyzed the effect of word position and stress on sound changes affecting unstressed vowels in Catalan using assimilatory and dissimilatory data collected by dialectologists and linguists since the beginning of the XX century. According to this database, progressive assimilations and dissimilations triggered by consonants on an immediately following unstressed vowel occur mostly in the word syllable-initial position (70.5% of the time) which is in support of the notion that segmental prominence plays an important role in sound change implementation. Table 2 (1a) exemplifies the assimilatory action of consonants placed in the absolute word-initial position: the raising processes /o/ > [u]

and /e/ > [i] triggered by preceding labial and alveolopalatal consonants, respectively, and the lowering process /e/ > [a] after the alveolar trill [r] whose production involves some tongue dorsum lowering and retraction.

The examples under (1b) in Table 2 show that vowel-dependent regressive assimilations are prone to apply when the triggering vowel is stressed and the target vowel is located word-initially. This finding is in agreement with stressed vowels being especially prominent while running against the assumption that unstressed vowels subject to a specific sound change should be located in a non-prominent word position. The relevant examples in the table include two vowel raising instances induced by a contextual high vowel (/o/ > [u], /e/ > [i]) and an instance of vowel lowering triggered by a contextual low vowel (/e/ > [a]).

The third set of examples reported in section (1c) of Table 2 is about changes in unstressed vowel quality which are exerted simultaneously by contextual vowels and consonants. While sound changes have been traditionally attributed to a single contextual phonetic segment, there are cases where several contextual segments may act as sound change triggers.

Table 2 - *Assimilatory changes in unstressed vowels (underlined) triggered by contextual vowels and consonants (in bold). Data have been taken from a Catalan database*

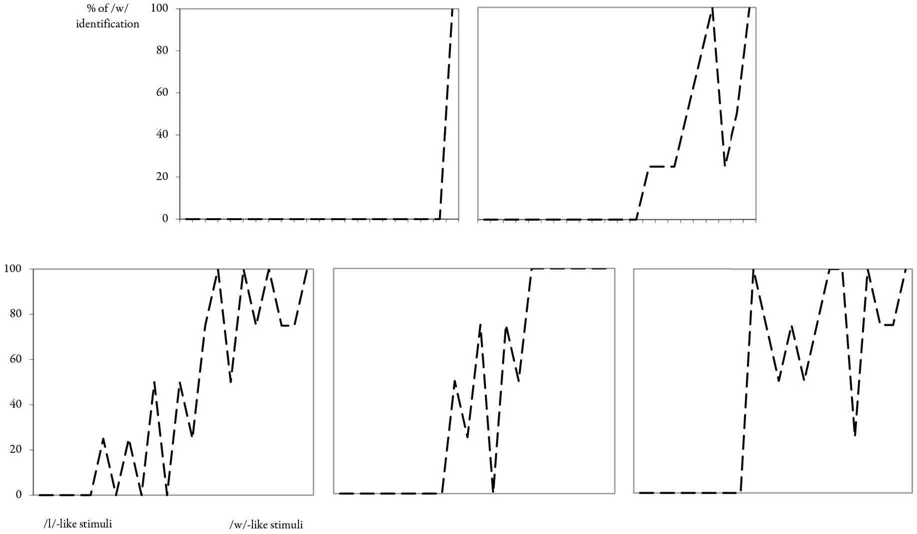
Process	Triggering segment	Examples
(1a) Progressive assimilations induced by word initial consonants		
/o/ > [u]	labial C	[**b**oˈro] “bud”
/e/ > [i]	alveolopalatal C	[**ʎ**esˈto] “lath”
/e/ > [a]	alveolar trill	[**r**eˈse] “shelter”
(1b) Regressive assimilations induced by vowels		
/o/ > [u]	high V	[toˈs**u**t] “stubborn”
/e/ > [i]	high V	[reˈz**i**na] “resin”
/e/ > [a]	lower V	[teˈr**a**t] “terrace”
(1c) Assimilations induced by contextual vowels and consonants		
/o/ > [u]	labial C, alveolopalatal C, high V	[**p**oˈ**ʎi**] “donkey”
	velar C, labial C, high V	[**k**oˈ**mu**] “common”
/e/ > [i]	alveolopalatal C, high vowel	[**ʎ**eˈɣ**u**m] “vegetable”
		[deˈ**ʒu**] “on an empty stomach”

Thus, for example, the raising of /o/ to [u] in the word [poˈʎi] “donkey” in Western Catalan may in principle be exerted at the same time by the preceding labial consonant (which causes an increase in degree of labiality during /o/), and the following alveolopalatal consonant and high vowel (which cause the tongue dorsum to raise). Three contextual segments may also be involved in the raising of /o/ in the word [koˈmu] “common” in the same Catalan dialect: the stressed high back vowel and the following labial consonant, as well as the preceding velar stop which may cause the postdorsal closure for /o/ to shift from the upper pharynx towards the velar zone. Finally, the change /e/ > [i] in the words [ʎeˈɣum] “vegetable” and [deˈʒu] "on an empty stomach" also in Western Catalan may be induced by both the following high stressed vowel and the preceding or following alveolopalatal or palatoalveolar consonant. All these examples reveal that the regressive assimilatory effect of a distant vowel may be reinforced by the progressive or regressive action of a contextual consonant, and therefore that vowel assimilations may be induced by the coarticulatory influence of several contextual segments acting simultaneously.

4. *How does sound change originate in the individual?*

Ohala's hypocorrection model states that sound change originates from the failure to compensate for coarticulation, which implies that listeners with little knowledge of the language such as children and second language learners are most likely to make perceptual errors and initiate sound changes. The lack of communication among peoples speaking the same language may cause much dialectal diversity to occur and therefore is also the seed for sound change. It has been shown indeed that individuals living in islands or remote regions may exhibit phonetic developments which diverge from the rule (Andersen, 1998). Another factor contributing to the inception of sound change, which is of much concern in the present investigation, is the degree of sensitivity on the part of the listener to small phonetic differences in the realization of a given phoneme which may take place in different contextual, positional and prosodic conditions. In so far as they are highly sensitive to such small variations in the acoustic signal, innovative listeners may lead sound changes (Grosvald, Corina, 2010). The relevance of this latter aspect can be ascertained from results obtained from peceptual tests which require a fine discrimination between acoustic stimuli differing slightly in one or more acoustic cues. Thus, the 17 subjects who took the identification test for preconsonantal [ɫ] referred to in section 1.2.2 could be grouped depending on their degree of sensitivity to small spectral differences which were available in the acoustic stimuli: at one end of the sensitivity scale, five subjects yielded no /w/ identification percentages at all and another two subjects seldom heard /w/ (two top panels in Figure 2); at the other end, ten subjects heard more instances of /w/ as the [ɫ] spectral cues became more /w/-like, with seven of them showing a more gradual increase in /w/ identification percentages than the remaining three (three bottom panels in Figure 2).

Figure 2 - *Subject-dependent patterns of identification of dark /l/ as /w/ according to the perception test described in section 1.2.2 (Top) Poor /w/ identification patterns; (bottom) better identification patterns exhibiting gradual and abrupt rises in the number of /w/ responses. In all graphs, the stimuli placed on the left are /l/-like and those on the right are /w/-like*



5. *Summary and conclusions*

Production and perception data reported in this paper reveal that intermediate articulatory variants need to be looked for before disclaiming an articulation-based explanation for specific sound changes. The relevance of phonetic detail for uncovering the phonetic causes of sound change has been illustrated in the case of two changes which are prone to be articulation-based rather than triggered by acoustic equivalence: /l/ vocalization and velar softening. Along these lines the study has also shown for /l/ vocalization and the segmental decomposition of palatal consonants that sound changes may have more than one acoustic origin and therefore may proceed through two and perhaps more evolutionary pathways. Information about the acoustic cues which contribute to the misidentification of one phoneme by another in these circumstances may be obtained by manipulating fine-grained acoustic characteristics which correlate with articulatory variations that may take place in real speech scenarios.

Results reported in this study also suggest that the acoustic cues for a given phoneme may play a primary or secondary role in sound change inception depending on their relative salience, which may vary from dialect to dialect. Secondary acoustic cues may be used by speakers to enhance the perceptual effectiveness of the primary acoustic cue. Thus, dark /l/ vocalization has been shown to depend on changes in F2 frequency at the steady-state consonant period and thus on tongue body configuration and apical contact loss rather than on variations in the timing of the vowel transitions which result from different degrees of gestural anticipation during the preceding vowel. On the other hand, velar softening appears to be implemented through a rise in burst intensity level rather than through burst lengthening. In the two cases the secondary acoustic characteristics may assist the primary cue in giving rise to a sound change.

The role of segmental prominence and the individual speaker has also been looked into. As for the former aspect, descriptive data on the frequency of occurrence of contextual segments which may trigger a given change suggest that prominence within the word domain plays a relevant role in the implementation of assimilatory changes affecting unstressed vowels. Moreover, it appears that sound changes may be triggered by more than one segment at a time, a possibility which deserves being explored by means of acoustic analysis and perceptual tests in the future. Regarding the role of the individual speaker, the paper argues that those individuals who are especially sensitive to small variations in the acoustic signal are prone to lead sound changes.

Acknowledgments

This research has been funded by projects FFI2013-40579-P from the Spanish Ministry of Economy and Competitiveness, by ICREA (Catalan Institution for Research and Advanced Studies), and by the research groups 2014SGR61 from the Catalan government.

References

Andersen, H. (1988). Center and periphery: adoption, diffusion, and spread. In Fisiak, J. (Ed.), *Historical dialectology, regional and social*. Berlin: Mouton de Gruyter, 39-83.

Beddor, P.S. (1983). *Phonological and phonetic effects of nasalization on vowel height.* Bloomington: Indiana University Linguistics Club.

Carignan, C., Shosted, R.K., Shih, Ch. & Rong, P. (2013). Compensatory articulation in American English nasalized vowels. In *Journal of Phonetics*, 39, 668-682.

Coromines, J. (1990-1998). *Diccionari etimològici complementari de la llengua catalana,* 8 vols. Barcelona: Curial.

Fougeron, C., Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. In *Journal of the Acoustical Society of America*, 101, 3728-3740.

Garrett, A., Johnson, K. (2010). Phonetic bias in sound change. In Yu, A.C. (Ed.), *Origins of sound change: Approaches to phonologization*. Oxford: Oxford University Press, 51-79.

Gauchat, L. (1925). Confusions d'occlusives dans les patois de la Suisse romande. In *Homenaje ofrecido a Menéndez Pidal*, 1, Madrid: Hernando.

Grammont, M. (1933). *Traité de phonétique*. Paris: Delagrave.

Grosvald, M., Corina, D. (2010). The production and perception of sub-phonemic vowel contrasts and the role of the listener in sound change. In Solé, M.J., Recasens, D.

(Eds.), *The Initiation of sound shange: Perception, production, and social factors*. Amsterdam: John Benjamins, 77-102.

Guion, S.G. (1998). The role of perception in the sound change of velar palatalization. In *Phonetica*, 55, 18-52.

Hall, R.A. (1974). *External history of the Romance languages*. New York: Elsevier.

Hardcastle, W.J., Barry, W.J. (1985). Articulatory and perceptual factors in /l/ vocalization in English. In *University of Reading Phonetics Laboratory Work in Progress*, 5, 31-44.

Hickley, R. (1984). On the nature of labial velar shift. In *Journal of Phonetics*, 12, 345-354.

Labov, W. (1994). *Principles of linguistic change: Internal factors*. Wiley-Blackwell.

Lausberg, H. (1965). *Lingüística románica*, 1, Madrid: Gredos.

Lin, S., Beddor, P.S. & Coetzee, A.W. (2014). Gestural reduction, lexical frequency, and sound change: a study of /l/ lenition. In *Laboratory Phonology*, 5, 2-36.

Lisker, L. (1986). "Voicing" in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. In *Language and Speech*, 29, 3-11.

Menéndez Pidal, R. (1968). *Manual de gramática histórica española*. Madrid: Espasa Calpe.

Millardet, G. (1910). *Études de dialectologie landaise; le développement des phonèmes additionnels*. Toulouse: E. Privat.

Ohala, J.J. (1974). Phonetic explanation in phonology. In Bruck, A., Fox, R. & La Galy, M. (Eds.), *Papers from the parasession on Natural Phonology*. Chicago: Chicago Linguistic Society, 251-274.

Ohala, J.J. (1978). The story of [w]: An exercise in the phonetic explanation for sound patterns. In *Report of the Phonology Laboratory*, University of California at Berkeley, 2, 133-155.

Ohala, J.J. (1981). The listener as a source of sound change. In Masek, C.S., Hendricks, .A. & Miller, M.F. (Eds.), *Papers from the Parasesssion on Language and Behaviour*. Chicago: Chicago Linguistic Society, 178-203.

Ohala, J.J. (1989). Sound change is drawn from a pool of synchronic variation. In Breivik, L.E., Jahr, E.H. (Eds.), *Language change: Contribution to the study of its causes*. Berlin: Mouton de Gruyter, 173-198.

Phillips, B.S. (2006). *Word frequency and lexical diffusion*. New York: Palgrave Macmillan.

Rankin, R.L. (1974). Latin kw, gw Rumanian p,b: an explanation. In Luján, M., Hensey, F. (Eds.), *Current studies in Romance linguistics.* Washington: Georgetown University Press, 14-26.

Recasens, D. (2014). *Coarticulation and sound change in Romance*. Amsterdam: John Benjamins.

Recasens, D. (2011). Velar and dental stop consonant softening in Romance. In *Diachronica*, 28, 186-224.

Recasens, D. (2014). Acoustic and coarticulatory characteristics of (alveolo) palatal stop consonants, and velar softening. In *Journal of Phonetics*, 42, 37-51.

Recasens, D., Espinosa, A. (2010). A perceptual analysis of the articulatory and acoustic factors triggering dark /l/ vocalization. In Recasens, D., Sánchez Miret, F. & Wireback, J. (Eds.), *Experimental phonetics and sound change*. München: Lincom Europa, 71-82.

Recasens, D., Espinosa, A. (2010). The role of spectral and temporal cues in consonantal vocalization and glide insertion. In *Phonetica*, 67, 1-24.

Repetti, L., Tuttle, E.F. (1987). The evolution of Latin PL, BL, and CL and GL in Western Romance. In *Studi mediolatini e volgari*, 33, 53-115.

Rohlfs, G. (1966). *Grammatica storica della lingua italian e dei suoi dialetti*. 1. Fonetica. Turin: Einaudi.

Rousselot, J.-P. (1897-1901). *Principes de phonétiques expérimentale*. Paris: Welter.

Scobbie, J., Wrench, A. (2003). An articulatory investigation of word final /l/ and /l/-sandhi in three dialects of English. In Solé, M.J., Recasens, D. & Romero, J. (Eds.), *Proceedings of the 15th International Congress of the Phonetic Sciences*. Barcelona: Causal, 1871-1874.

Solé, M.J., Recasens, D. (2012). *The initiation of sound change. Production, perception, and social factors*. Amsterdam: John Benjamins.

Thomsen, V.L.P. (1876). Remarques sur la phonétique romane. L'<i> parasite et les consonnes mouillées en français. In *Mémoires de la Société de Linguistique de Paris*, 3, 106-123.

Wang, W.S.-Y. (1969). Competing changes as a cause of residue. In *Language*, 45, 9-25.

Yu, A. (2013). *Origins of sound change. Approaches to phonologization*. Oxford: Oxford University Press.