Anatomy of the Korean Plosives

Dick Grune
dick@dickgrune.com

Febr. 2016; DRAFT; needs updating

Abstract

The observed characteristics of the Korean plosives (stops) system, including the assimilation rules, are explained on the basis of a simple generative model.

Introduction

A short inspection of the Internet shows that there are many pages and papers about the nature of the Korean plosive (stop) triads p, ph, pp; t, th, tt; k, kh, kk; and j, ch, jj.[1] The reason is that the Standard Model (based on properties like 'voiced', 'aspirated', 'glottalized', etc.) cannot give a satisfactory description of them.

This paper shows that the properties of the Korean plosives can be explained by a simple model with a single parameter.

1 The Korean stops system

Each Korean plosive C occurs in three forms, distinguished both in pronunciation and in writing: C, C^h, CC. They are traditionally called 'lax', 'aspirated', and 'tense', respectively. Each of these is pronounced differently at the beginning of a word ('initially'), between two vowels ('intervocally'), at the end ('finally'). This results in nine combinations. Assimilations rules apply when two consonants come together; this occurs only between vowels. Assimilation rules are covered in Section 4.

From accounts of native speakers (for example Sohn (1994)[2]), many Internet pages (for example, Wikipedia), and personal observation, the three forms of plosives can be described as follows (to avoid the abstract plosive C, we shall use the bilabial plosives p, p^h, and pp in the examples):

[1] The j, ch, jj may be perceived as affricates by Westerners, but Sohn, pg. 434, writes: "The dorso-palatal consonants /c, ch, c'/ are regarded as plosives, and not as affricates, in that each of the phonemes is produced and perceived as a single sound, and not as the coarticulation of a plosive and a fricative.", and indeed no special rules apply to it in Korean morphology.

[2] Ho-min Sohn, *Korean*, Routlegde (1994), Chapter 3

- lax
  - initial

    A soft, voiceless stop, followed by a light aspiration. It is not like the Chinese unmarked b as in 'Beijing', in that it is clearly voiceless, nor the Danish devoiced b, [b̥], in billig, in that it is more precise. The aspiration is much lighter than that in an English or German p, more like a following soft h. So [ph] is probably a good description.
  - intervocal

    Usually a voiced stop, similar to English, German or Dutch [b]. But often a voiced fricative is heard, f.e. [jɔβoseo] for 여보세요 'hello' or [mɔɣɔjo] for 먹어요 'eat'.
  - final

    A voiceless stop with no audible release: [p̚].
- aspirated
  - initial

    A voiceless aspirated stop, stronger than a lax one. The aspiration is more German than English, in that the release of the stop is not slower than in the unaspirated version; in particular there is no trace of the t sound acquiring a slight s ending, as it does, for example, in the English word type which is often pronounced as [tˢajp]. A voiceless stop, followed by the strong Korean h, is probably a good description: pɦ.
  - intervocal

    A voiceless stop with aspiration of variable length. The student listening tracks contain many examples where the aspiration is audible, but also a large minority where there is definitely no aspiration, and only a voiceless stop remains.
  - final

    A voiceless stop with no audible release: [p̚].
- tense
  - initial

    Almost always a forceful voiceless stop, with no trace of aspiration. Remarkably the student listening tracks contain some examples of tense initial stops that I would not hesitate to call voiced.
  - intervocal

    A long voiceless stop, very much like the pp in Italian appare 'he appears'.
  - final

    A voiceless stop with no audible release: [p̚].

Syllables starting with a tense stop are usually louder and of higher pitch than those starting with other consonants. The same is true to a lesser extent for aspirated stops: cf. 안에 (pronounced 아네, [ane]) 'in' with even stress and 앞에 (pronounced 아페, [apɦe] 'in front of' with slightly stronger and higher e.

2 The hypotheses

A problem with the above description is that it is couched in relative and subjective terms: soft, stronger than, faint, etc. By contrast, the Standard Model deals in absolutes: a consonant is voiced or not, aspirated or not, etc. However it can be pointed out that the Standard Model is already inadequate for the Dutch fricative pairs `f`/`v` and `s`/`z`. Each of these pairs consist of a voiceless and a half-voiced member; a half-voiced fricative starts voiceless and ends voiced. On the basis of this distinction words like `fier` 'proud' and `vier` 'four', and `saai` 'dull' and `zaai` 'sow seed' can clearly be distinguished in a large part of the Netherlands.[3]

Rather than attempting to give a different descriptive classification, we propose a generative model that leads to the phonemes described above. It is based on two simple hypotheses:

1. Air flow to the speech apparatus is continuous and of constant speed.

2. The members of the triads differ in one parameter only: the force with which the obstruction that forms the stop is held closed. This parameter has three possible values, very low, normal, and very high.

When the closing of the airways occurs at the start of the consonant, according to Hypothesis 1 the air pressure builds up. When the pressure reaches, or slightly exceeds, the force with which the obstruction is kept shut, the obstruction opens and the sound of the consonant begins. Exactly *when* that happens and *what* happens next, depends on that the magnitude of that force.

All properties described in Section 1, the assimilation rules and a few other characteristics derive directly from these hypotheses, as we will show below.

The above statements are hypotheses; I have neither the equipment nor the expertise to prove them physically, but they presently derive their justification by their explanatory power. For related measurements see Broersma (2010)[4].

3 Application of the hypotheses

We will first demonstrate the mechanism based on the above hypotheses by showing how the tense initial stops are generated; in these stops the air flow is stopped with considerable force. We will then use this mechanism to explain the formation of the simple stops (Section **??**), and the `jieut` triad, which does not sound like a stop but rather like an affricate (Section **??**). The assimilation rules and the nasal consonants will be treated in Sections 4 and 5. The `siot` pair, which is usually mentioned together with the above triads, is not a stop at all, and will not be treated here; see, however, the Conclusion (Section 6).

We will use variants of the following diagram in the explanations:

[3] Along the west coast, both `v` and `z` are voiceless, while in the south they are both fully voiced.

[4] Mirjam Broersma, *Korean lenis, fortis, and aspirated stops: Effect of place of articulation on acoustic realization*, InterSpeech 2010, pp. 941-944.

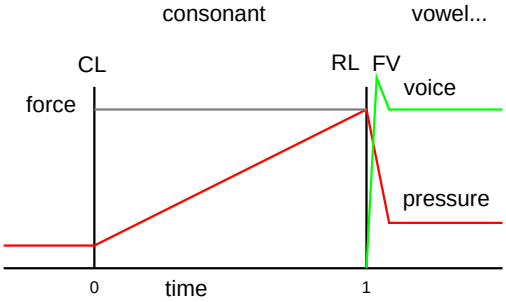

Figure 1. Demonstration of the diagram

Time runs from left to right. The vertical line marked `CL` at time 0 represents the closure of the stop position (lips for `p`, tongue against upper gum ridge for `t`, and back of the mouth for `k`). The vertical line marked `RL` at time 1 represents the release of the stop. There is also a dashed line at time 1, obscured here by the `RL` line, which is marked `FV` for Fully formed Vowel. The grey line shows the force with which the closure is kept shut; the red line shows the air pressure as a function of time; and the green line shows the strength of the voice.

We now follow the time line. Before time 0 there is no constriction, so the pressure is low. At time 0 the airway is closed with a strength indicated by the level of the grey line, and since by Hypothesis 1 the air keeps coming, the pressure builds up. At time 1 the pressure meets the force of the closure, the closure is pressed open, and the release immediately causes a strong onset of the vowel. After the release the pressure drops, but not to the level before time 0; the excess pressure is used to fuel the voicing of the vowel. The result is a "normal" long voiceless stop, as described in Section 1.

The lines for force, pressure, etc., are schematic; in reality the lines will have more naturally curved shapes.

We will now consider the various stops in their three positions, initial, intervocal, and final.

3.1 The 'lax' stops

As the name already implies, the force holding the closure of a lax stop shut is very weak. For an initial lax stop this means that upon closure the pressure begins to raise linearly, but little pressure is needed to open the closure, so the release moment is very soon after the closure moment, and since the pressure difference is low, a soft plosive results. Another consequence of the low pressure at release is that some time is needed to reach full vowel strength. During this time aspiration is heard, as indicated by the brown line:

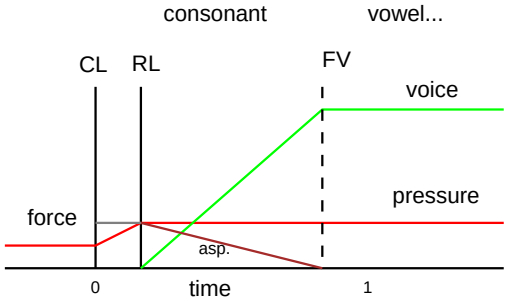

Figure 2. Lax initial

Intervocally, the closure and release are too close together to interrupt the voicing of the vowels, and a "normal" voiced stop results:

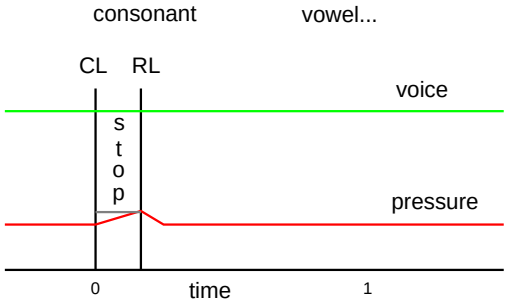

Figure 3. Lax intervocal

In many cases the closing force associated with the lax consonant is not even enough to stop the vowel from the previous syllable, no stop occurs, and a fricative is heard, in accordance with Section 1.

Finally, i.e. at the end of a word, the closure of the stop is audible because the vowel ends, but the release is inaudible. The release, which must eventually follow, is indicated as a dotted line in the diagram:

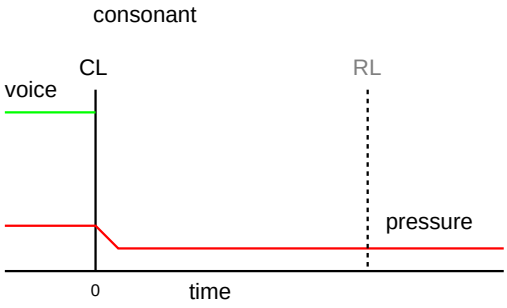

Figure 4. Final (lax, aspirated, and tense)

Since the members of a triad differ only in the closing force at the moment of their generation and since that force is immaterial at the end of a word, it follows that all three members of a triad sound the same in final position.

3.2 The ‘aspirated’ stops

The aspirated stop is characterized by much stronger closure force resulting in a release moment much further on along the time line:

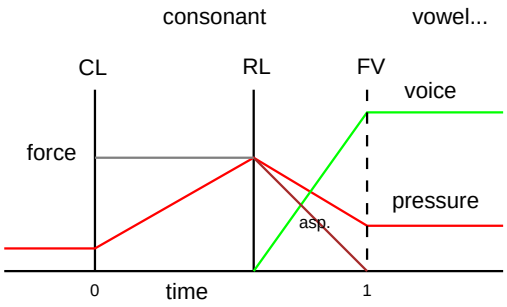

Figure 5. Aspirated initial

Essentially the same happens as in the lax stop, but due to the longer duration much more pressure is built up, resulting in a much larger surplus of air, i.e. aspiration: the area under the brown line marked `asp.` is much larger than in Figure 2.

The intervocal version of the aspirated stop is almost the same as the initial one:

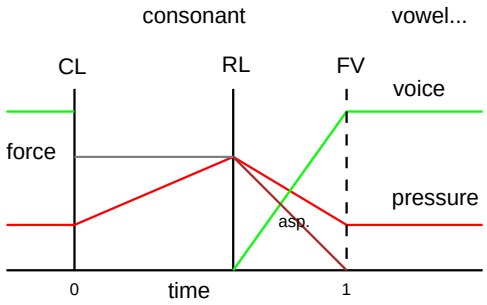

Figure 6. Aspirated intervocal

The gap is too large for the voice to keep going, so it stops and has to be restarted. This results in a sound very similar to that of the initial aspirate.

The frequent occurrence of an (almost) unaspirated intervocal aspirate may be explained if we assume that restarting the voicing is easier intervocally due to the voice that has just stopped, and therefore occurs earlier:

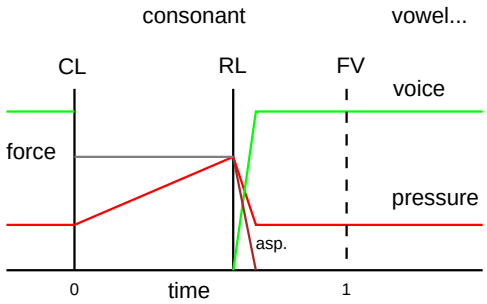

Figure 7. Aspirated intervocal, with reduced aspiration

3.3 The 'tense' stops

The generation of the initial tense stop has already been explained at the beginning of this section (Figure 1).

It has, however, been observed that the vowel after a tense stop is often louder than after other stops. So it seems that the spike level due to the high pressure release shown in Figure 1 continues throughout the whole vowel:

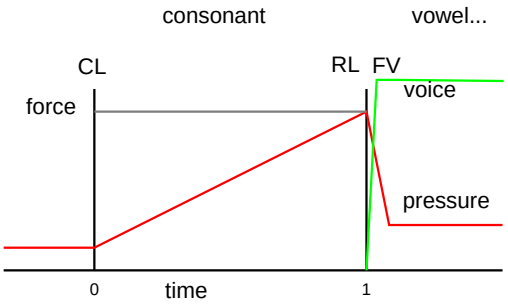

Figure 8. Tense initial

The diagram for the intervocal tense stops is almost identical:

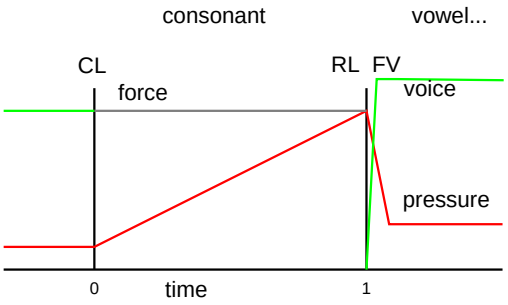

Figure 9. Tense intervocal

Comparison of the diagrams for the three types shows that those for "tense" are simpler than the others. This agrees with the observation that tense stops are the first to be learned properly

by Korean children, and with the explanation which Kong, Beckman, and Edwards (2011)[5] give for that phenomenon.

4 The assimilation rules

There are six more consonants besides the stops: the three nasals `m`, `n`, and `ng`; and a rest class consisting of `s/ss`, `r/l` and `h`. Each of these, and the stops themselves, can end up next to a stop, and when that happens some form of assimilation may occur. Some assimilation rules are easy to understand, for example when a stop meets a `h`, the stop becomes aspirated: 좋다 `joh-da` 'to be good' is pronounced `[jotha]`; some are more opaque, for example those for combining the `r/l` with the nasals; but the two most important ones, stop-stop and stop-nasal, can be easily understood using the model presented here.

4.1 Reinforcement

For two stops to come together, one must be at the end of the first syllable and the other at the beginning of the second syllable, as for example in 잡다 `japda` 'to catch'. As is normal at the end of a syllable, the `p` closes, and does not open again; next the stop pair is given the full length of a consonant, during which the pressure rises; finally the `d` opens:

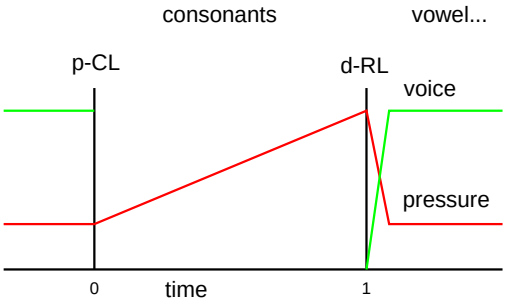

Figure 10. Reinforcement: `pd` → `p͡tt`

This diagram looks strikingly like Figure 9, the tense intervocal; the only difference is that the closure and release occur at different articulation points. Indeed the combination `pd` is pronounced and heard as `p͡tt`: so we have `pd` → `p͡tt`. If both stops are the same the closure and release occur at the same articulation point, and the picture becomes identical to Figure 9; indeed 학교 `hak.kyo` 'school' is pronounced `[ha.kkyo]`. These are examples of "reinforcement rules".

In these examples both stops are lax, but the same happens when the first stop is aspirated or tense, or if the second stop is tense: 붙다 `puth-da` 'to stick' is pronounced `[put͡.tta]`; 닦다 `dakk-da` 'to brush' is pronounced `[dak͡.tta]`; and 닭띠 `dalk.tti` 'The Year of the Chicken"[6]is pronounced `[dak͡.tti]`.

[5]Eun Jong Kong, Mary E. Beckman, and Jan Edwards, *Why are Korean tense stops acquired so early: The role of acoustic properties*, J. Phon. 2011, 39(2): 196-211

[6]Words with a tense second stop are rare.

The situation is different, however, when the second stop is aspirated. The air flow is stopped at the position of the first stop, and then continued at the position of the second stop, at the normal time for aspirated stops. This means that 식탁 `sik.thak` 'dinner table' is just pronounced `[sik͡.thak͡]`.

4.2 Nasalisation

When is stop is followed by a nasal, as for example in 합니다 `hapnida` 'does', the `p` closes, and is given the full length of a consonant, just as above, but rather than stopping the voice and creating a short hiatus, the voice switches to nasal (blue line), in anticipation of that of the following `n`:

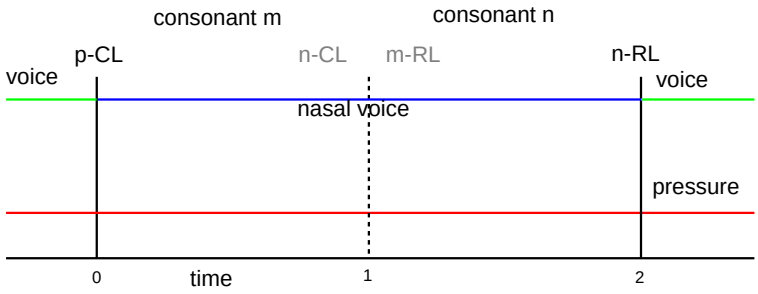

Figure 11. Nasalisation: `pn` → `mn`

Somewhere between time 0 and time 2 (the end of the `n`) the switch-over from the closure of the lips to that of the tongue is negotiated; it must occur in the order indicated in the diagram, or the nasality would be broken. At time 2 the tongue is released; this ends the nasality of the voice, which returns to normal voice, ending the `n`. So underlying `hapnida` is pronounced `hamnida`.

The whole process does not require a change in pressure, which may explain why Koreans claim that [mn] is much easier to pronounce than [pn], an evaluation not always shared by West-Europeans.

5 Initial `n` sounding like `d`

Although not mentioned in any text book I know, it is widely observed on the Internet[7] and in the classroom that the Korean initial `n` often sounds like a full-voiced `d`: 네 `ne` 'yes' is heard as `de`, 누구 `nugu` 'who' is heard as `ndugu` or `dugu`. Something similar can be observed with the initial `m`: 무엇 `muot` 'what' is heard as `buot`, 마당 `madang` 'yard' is heard as `badang`. Native speakers, however, swear that they say just `n` and `m`.

The phenomenon can be explained simply in the present framework by first realizing that nasal consonants and vowels are formed by lowering the velum so air can escape through the nose. Pronouncing an `n` followed by a vowel then requires the raising of the velum and the release of the tongue to be performed at exactly the same moment (at time 1):

[7] for example Kim, Young Shin – An Acoustic, Aerodynamic and Perceptual Investigation of Word-Initial Denasalization in Korean, PhD Thesis, University College London, 2011

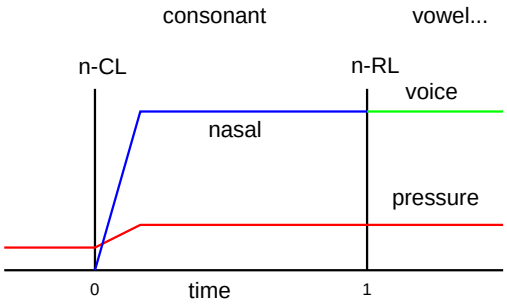

Figure 12. An initial n

Now suppose this synchronization is not perfect. There are two ways in which this can occur: release the tongue and then raise the velum, in which case the vowel becomes nasalized; or raise the velum and then release the tongue. In that case we get:

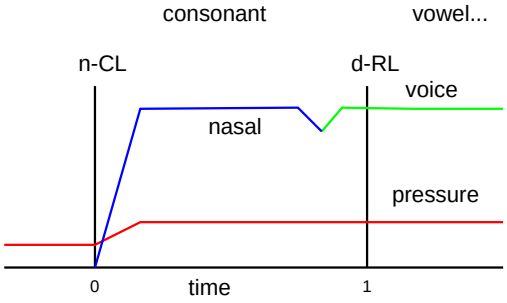

Figure 13. The initial n turned into nd

The diagram shows that this produces a dental release during voicing, i.e. a d. Since there is no initial d in Korean its presence proves the presence of the n, so to reduce speech effort, the n can be whispered –be without voice– resulting in n̥d:

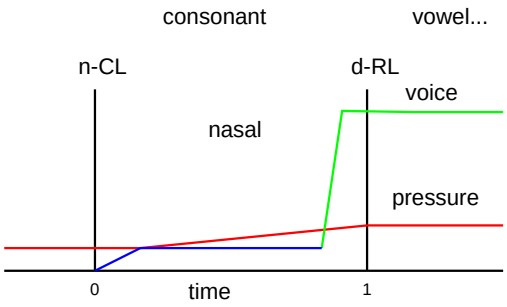

Figure 14. A whispered n followed by a d: n̥d

When the word that start with an n is closely preceded by a word ending in a vowel there is no need to start the n from 0 as in Figure 13, nor would whispering the n yield a speech effort reduction as in Figure 14, because the voice is already there.

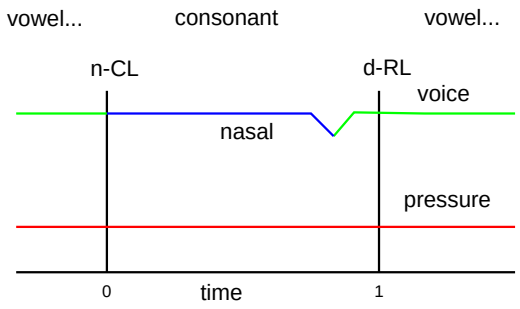

Figure 15. An initial `n` preceded by a vowel

As a result we often hear both the `n` and the `d` in such cases: 더 넓어요 `deo neolbeoyo` 'is wider' is heard as `deondeolbeoyo`.

This explains the observed facts, including native speakers' perception.

6 Conclusion

The properties of the Korean stops, including their assimilation rules, can be explained by a simple model based on constant air flow and uniform consonant length, with only one parameter: the time between closure and release of the stop.

Voice is seen to be a property of the vowels only, to be shared or not to the stops as dictated by the model.