

Towards a Dynamic View on L2 Phonological Acquisition

Adelaide H. P. Silva

Federal University of Paraná, Brazil
adelaidehpsilva@uol.com.br

1. Introduction

We argue in this paper that some data on the production of Brazilian students learning English as a foreign language provide evidence for a dynamical approach to explain some facts that lay beyond the acquisition of L2 phonological level. We focus specially on the so-called epenthesis facts, i.e., Brazilian students learning English as a foreign language are said to insert a final high front vowel in English words ending in obstruent consonants, since those consonants do not occur word-finally in Brazilian Portuguese.

However, an acoustic analysis of the production of some EFL students shows that their production is somehow close to the production of a native speaker, in the sense that the students are able to produce the same articulatory gestures involved in the native production. The difference between the native and the nonnative lies in the timing of the gestures. This observation implies that, contrary to what the literature usually claims, the learners do not insert or delete segments. So, traditional approaches¹ are not adequate to explain these data, because they propose that when learning a foreign language, students usually insert, delete or even change some features or entire segments. But it is clearly not the case of inserting, deleting or changing features or segments. Instead, it is the case of phasing, or tuning the gestures in the speech of the nonnative the same way the gestures are phased in the native speech.

It follows then that, if we consider the students already produce some sounds or some sequences of sounds they are usually said not to, the task of the teacher is not teaching the students new sounds, but perceiving how the students produce the target-sounds and teaching them how to phase the articulatory gestures to reach a production closest to the native's one.

2. Traditional accounts on L2 phonological level acquisition

When dealing with L2 phonological acquisition, one of the most current approaches claim that interphonology plays a special role during the acquisition process, since we can find in it facts from both languages – the subject's native language and the second language he is acquiring. It is also assumed that the phonology of the mother tongue can have some influence on the acquisition of the L2 phonology, so that the learners transfer to the second language facts from the phonology of their L1.

In this sense, when Brazilian Portuguese (BP) speakers learn English as a foreign language, they are said to transfer some facts from BP phonology into English. Let's take into account word-final codas: BP does not allow plosives and the fricatives /f,v/ in word-final position². English, on the other hand, is plenty of words containing these consonants in

¹ In this paper we call "traditional approaches" to the phonological models somehow inspired on Chomsky and Halle's generative phonology, as for instance Feature Geometry or Metrical Phonology. These are all theories of extrinsic timing. So, we take as synonyms the terms "traditional approaches" and "generative phonology". This is to oppose these models to more recent and innovative ones, the so-called "dynamic models of speech production".

² It is worth noting that loan words such as "ping-pong" or "picnic" had their orthography adapted to follow the phonotactic rules of Brazilian Portuguese, that changes them into trochaic dissyllabic words. (In the specific case

word-final codas. Based on a view such as the one that argues in favor of the interphonology and of the interference of L1 on L2 phonological acquisition, one possible hypothesis concerning the acquisition of English by BP native speakers is that they will not produce plosives word-finally – at least during initial learning levels. Moreover, the hypothesis can expect learners to adapt words containing plosives word-finally to BP phonotactics.

It follows then that, in the case of English words such as “pet”, “luck”, “cat”, “dad”, “Bob”, Brazilian Portuguese learners of EFL can be expected to produce an epenthetic high front vowel just after the plosives, thus turning these monosyllabic words into dissyllabic ones with the stress laying on the second syllable³ (from right to left). If we assume this hypothesis, we can also expect learners not to produce the morphophonological difference in pairs such as “pet/petty”; “luck/lucky”; “dad/daddy”. That is, we can expect learners to have a single production for both words in each pairs. Teacher’s task, then, is to make the students perceive that they do not produce the pairs because they insert a vowel in a place where it does not exist.

Our point in this paper is that this hypothesis we built above – and that is current in the literature dealing with EFL acquisition by Brazilian Portuguese speakers – derives from a generative perspective, i.e., a perspective based on Chomsky and Halle’s Generative Phonology. A generative perspective assumes that there are universal facts underlying the phonological or the syntactic level and that can be found in a number of languages, but there are also language-specific facts. These are the ones that interest us the most here, because in the phonological level facts such as syllable structure, stress and the phonotactics are said to be language-specific, i.e., each language selects the syllable structures it will use, as well as its stress pattern and its sounds inventory, the sequences of sounds and the way they will be arranged into the words, generating some sequences that are allowed for that language and preventing some sequences from occurring, since they are not allowed.

Teaching a second language, then, involves teaching the students these aspects that are language-specific, such as the phonotactics of L2. And, if the student does not produce some sounds or some sequences of sounds adequately, it must be the case of interference of L1 on L2. This interference can consist of insertion of segments – just like the case of the so-called epenthetic vowel in English words ending with plosives, we just mentioned above – as well as of deletion of segments or even the “change” of one feature or one segment into another one, that causes some “strange” production (this could be the case of Brazilian Portuguese learners who do not produce a dental fricative in English, but a labiodental one, instead).

The generative approach - adopted by authors such as Eckman (2004), Ellis (1994) or Flege (1995) – could be successful to deal with facts as the one we are focusing – the insertion of an epenthetic vowel in monosyllabic words ending in plosive consonants by Brazilian Portuguese speakers learning English as a foreign language – if students did in fact insert the vowel word-finally, as expected in the light of the hypothesis we sketched above.

In the next section we present some evidence that students indeed do not produce what they are said to produce. If this is so, a generative perspective is not adequate to deal with the facts we will show and, therefore, another theoretical perspective should be adopted in order to explain the facts occurring in the process of L2 phonological acquisition.

of “ping-pong” we have two trochaic dissyllables that form a compound word.) That’s why these words are written respectively as “pingue-pongue” and “piquenique” in Brazilian Portuguese.

³ According to Albano *et al.* (1999) words with stress on the second syllable are the most frequent in Brazilian Portuguese.

3. Looking at learners' speech: Acoustic data can reveal what they really produce

The data presented bellow are based on Ferreira's (2007) study on the production of the distinction in pairs such as "pet/petty"⁴, that took as subjects male and female adolescents who have been studying English for approximately 250, 300 at the time data were collected. These data were selected because they exhibit quite clearly the facts we want to discuss.

We must add that the experiment contained pairs such as the one mentioned inserted in the carrier sentence "Say _____, please." Besides the target words, the *corpus* had also distractors, in a format similar to the one of the target words. The data were all collected in an acoustically-treated chamber, at 44 kHz, and were then submitted to an acoustic analysis using the software Praat.

The productions of the Brazilian learners were all compared to the productions of a 16 year-old American girl (GA) who was a "control subject". She was born in Nashville (Tennessee) and has been living in Brazil for four months when data were collected. We turn now to an acoustic description of some of the Brazilian Portuguese speakers' data, in comparison to the American subject data. The first pair we look at is "luck/lucky".

3.1 How learners can produce the distinction "pet/petty"

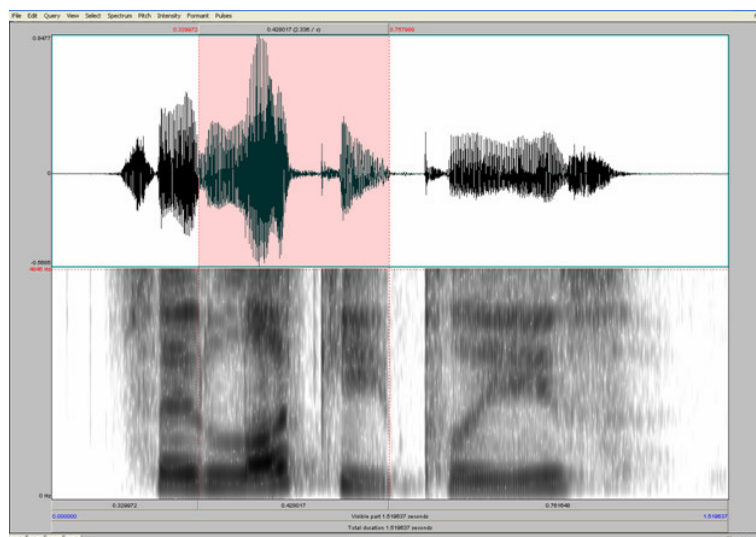


Figure 1. "Lucky" produced by the English native speaker (GA)

Figure 1 shows the waveform (upper) and spectrogram (lower) for the sentence "Say lucky, please", produced by the native speaker GA. The information in the waveform is aligned to the information in the spectrogram. The target word "lucky" stands between the vertical dashed lines in the spectrogram⁵. Let's turn now to the same word produced by one of the subjects of the experiment.

⁴ I was co-advisor of this study that makes an acoustic description of the production of 13 Brazilian Portuguese adolescent speakers learning English as a foreign language. The study focus on the distinction of pairs such as "pet/petty"; "bob/bobby"; "luck/lucky". Ferreira has kindly allowed me to use the set of data that I present in this paper.

⁵ This will be the same for all other figures in this paper, i.e., target words that are under examination stand all between vertical dashed lines.

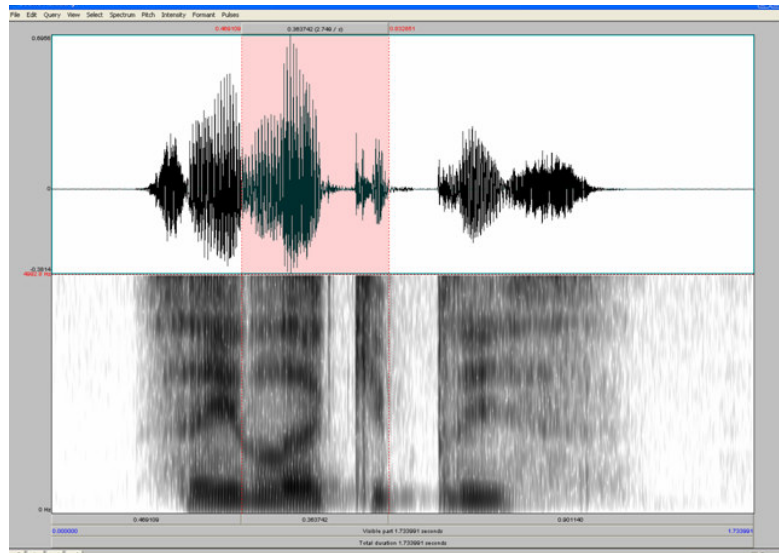


Figure 2. "Lucky" produced by the Brazilian Portuguese speaker (RF)

In comparison with the native speaker's production, we can say that the Brazilian Portuguese native speaker does produce a final vowel. But it is clearly shorter than the final vowel in the word "lucky" produced by the native speaker GA. Besides that, we can also notice that the first vowel of the word "lucky" is much longer in the speech of the Brazilian Portuguese speaker RF than in the speech of the English native speaker GA. It seems then that the difference between both speakers does not lie in the segmental level, but in the prosodic level, since in Brazilian Portuguese, according to authors such as Massini-Cagliari (1992) the acoustic correlates for word stress are intensity and also duration, i.e., stressed vowels have higher intensity than the others in the word and are also longer. RF seems to have produced a stress pattern very similar to the one found in trochaic disyllables in Brazilian Portuguese.

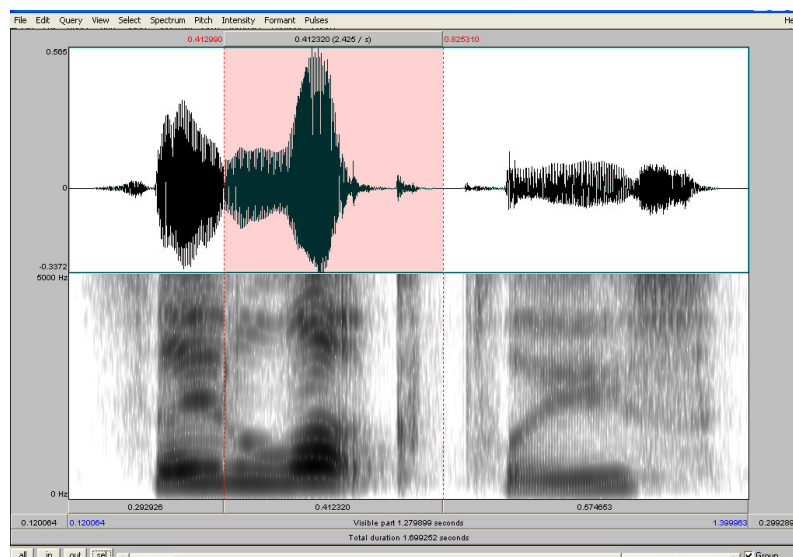


Figure 3. "Luck" produced by the native speaker GA

Figure 3, above, shows the production of the word "luck" by the native speaker GA. Notice, particularly, the glottal activity at the end of the word that causes noise to occur after the release for the velar plosive [k].

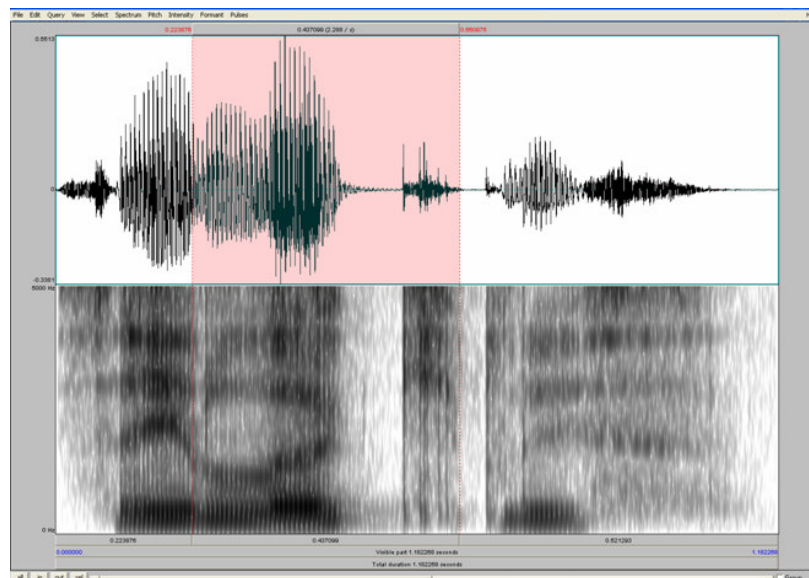


Figure 4. "Luck" produced by the Brazilian Portuguese speaker (RF)

Compared to the native speaker production in Figure 3, the production of the Brazilian Portuguese speaker in Figure 4 exhibits a great similarity, especially in word-final position: notice that, contrary to some expectations, the learner does not produce an epenthetic vowel after [k]. Moreover, we can see the same noise after the release for [k], due to glottal activity during the production of the plosive. Notice, however, that the noise is longer.

In the light of these data, then, we can say that Brazilian Portuguese speakers learning English as a foreign language can produce the distinction “luck/lucky”. It is not the case, then, that all learners have a single production for the two words in the pair. We can also hypothesize that in order to try to approximate his speech to the native’s, RF “exaggerated” in the glottal activity. This could be the reason why this acoustic event is clearly longer for RF than for the native speaker GA. Therefore we can say that we have here a first evidence to the fact that what happens in the acquisition of English as a foreign language by Brazilian Portuguese speakers, particularly in the case of the plosives in word-final codas, is that learners are able to produce the sequence of sounds that occur in the second language they are acquiring but do not occur in their mother tongue. It is just the case of adjusting the timing of the events relatively to the timing they have in the English language.

Let’s turn now to another pair of data, “pet/petty”.

3.2 How learners can produce the distinction “pet/petty”

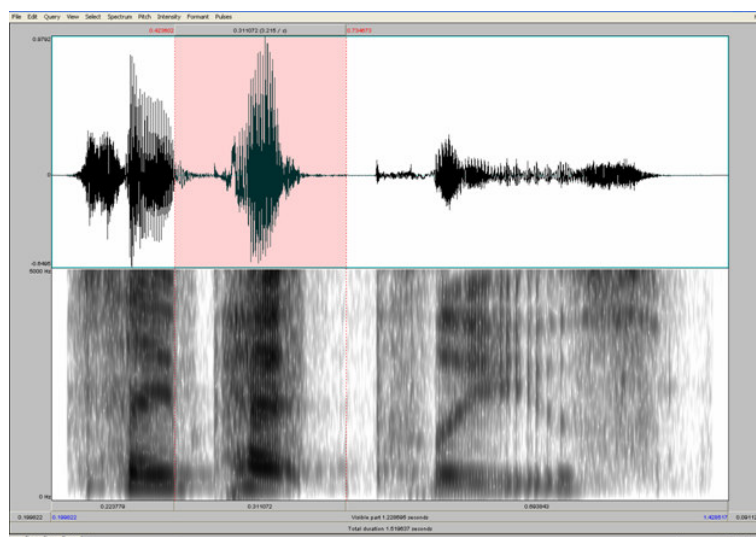


Figure 5. The word "pet" produced by the native speaker GA

A visual inspection of the speech signal, notably during the word “pet”, allows us to notice that, as expected, after the burst of the plosive [p] there is noise that turns the consonant to be aspirated. After the alveolar plosive [t] there is also some noise, probably caused by the position of the tongue when releasing the closure for [t], i.e., just after the tongue touches the alveolar region to produce the plosive, it is released and as a result it makes a constriction very similar to the one required by the homorganic fricative. It is then the case that we have a little frication after [t]. This seems to be a mechanical fact, related to the position the tongue has to assume in order to produce final [t] and the position it has to assume immediately after it, to produce the plosive [p], that is the next sound in the speech chain. Let’s turn now to the production of the same target word by a Brazilian Portuguese native speaker.

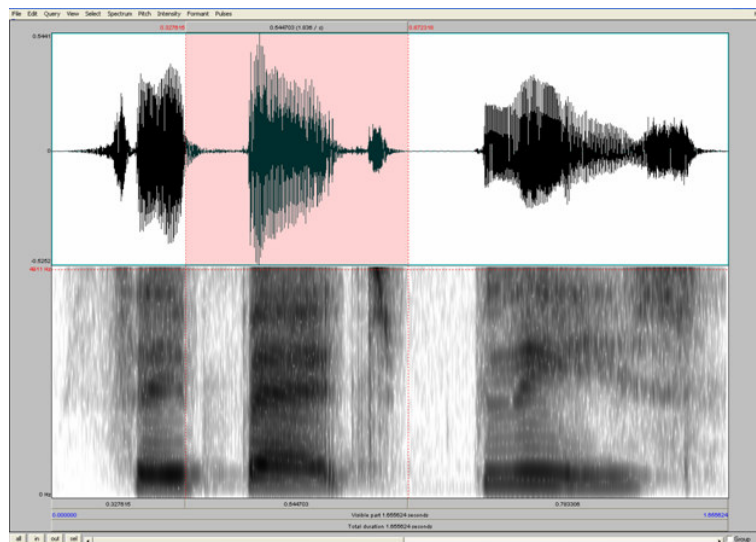


Figure 6. The word "pet" produced by the Brazilian Portuguese native speaker RH

First of all it is important to notice that we refer now to data of another Brazilian Portuguese native speaker, RH – remember that in the last section we referred to data of the subject RA, when dealing with the pair “luck/lucky”. The option of presenting data obtained from another subject is related to the aim of showing that the similarities of production strategies between a native speaker and a nonnative one is not restricted to one single learner, but can be found in data obtained from other students.

And, before making a visual inspection of the signal in the spectrogram of Figure 6, we also have to say that most Brazilian Portuguese dialects exhibit the frication of the alveolar plosives [t,d] when they precede the high front vowels [I, i], i.e., the plosives [t,d] are produced as their affricate counterparts in the vocalic context just mentioned. It can also be the case that in some specific contexts – e.g., word-finally, like in words like “leite” (milk) - the affricate is produced without the following high front vowel [I]. That’s probably because the “color” of the vowel can be found in the consonant, since both the affricate consonant and the vowel are produced in the same place of articulation. We can then expect Brazilian Portuguese native speakers to produce a word like “pet” with an affricate, instead of a plosive in the word-final coda, but not necessarily followed by a high front vowel.

Looking at the spectrogram for “pet” in Figure 6, above, we notice that indeed the Brazilian Portuguese speaker, learning English as a foreign language, produced a noise after the burst for [t], i.e., he produced an affricate consonant. Comparing the facts in this Figure 6 with the ones in Figure 5, we can see some similarity between them: obviously the Brazilian Portuguese speaker does not produce an aspirated [p] word-initially⁶. But, in the word-final coda, both speakers produce a noise, a frication noise. Just as in the case of the production of the word “luck”, the difference between the native speaker and the nonnative one lies on the duration of the acoustic event that occurs word-finally. In that case the effect was the glottal activity; in this case, it is the frication noise. Maybe it’s the case that nonnative speakers have to have a longer event in order to perceive their production as closer to a native speaker production.

We observe now some productions for the word “petty”.

First of all, it is important to note that the native speaker varies her production. We had three repetitions of the experimental *corpus* and, in each of the three, GA had different production strategies.

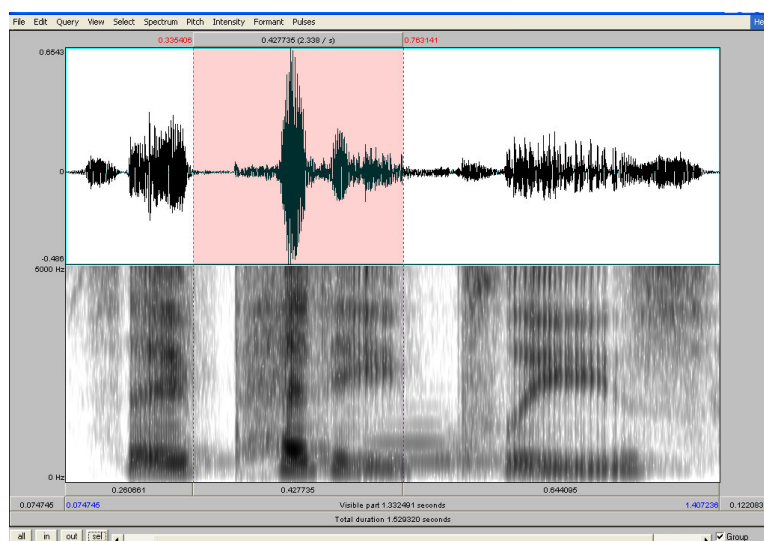


Figure 7a. The word “petty” produced by the English native speaker GA.

⁶ Portuguese does not have aspirate plosives.

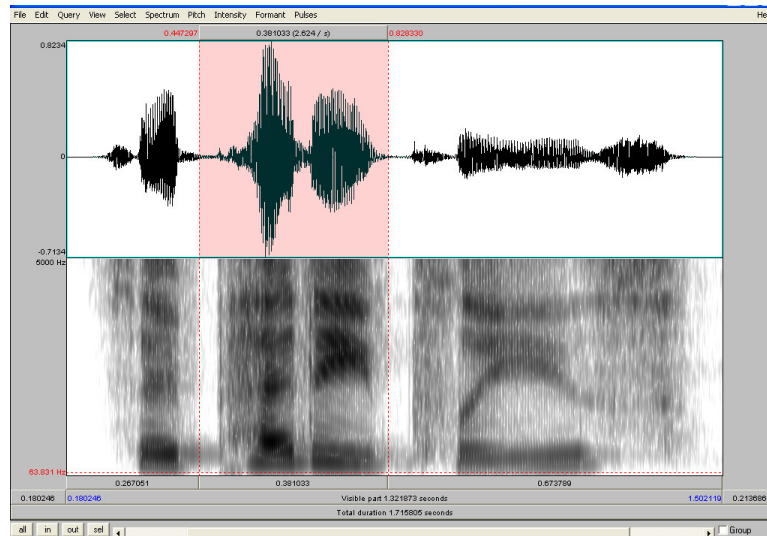


Figure 7b. The word “petty” produced by the English native speaker GA

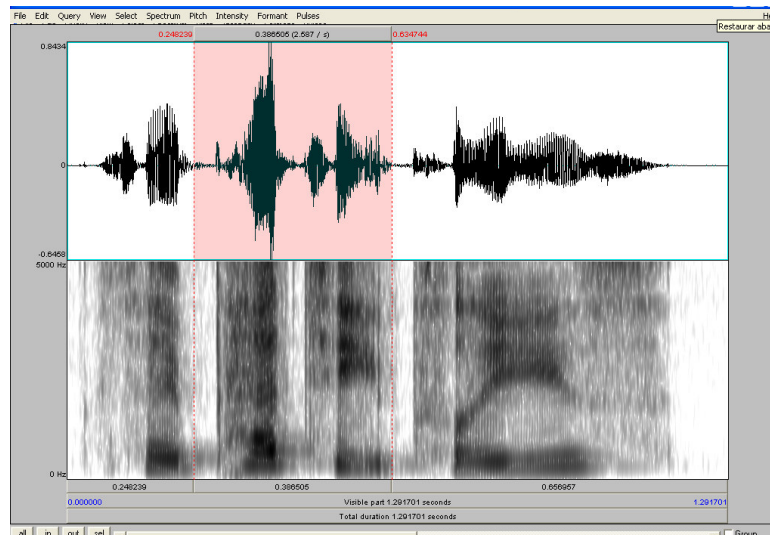


Figure 7c. The word “petty” produced by the English native speaker GA

The three productions for the word “petty” show that flapping is not a categorical rule, as already pointed out by Fox and Terbeek in a 1977 paper. Notice that flapping is indeed present in Figure 7b. In Figure 7c, on the other hand, flapping was not applied, so that we can see an alveolar plosive between the two vowels. Observe also that after the burst for [t] there is a frication noise. Finally, Figure 7a shows something that seems neither flapping properly nor an alveolar plosive. It seems to be an intermediate production between the one in Figure 7b and the other in Figure 7c. Anyway, roughly speaking, we can say that the native speaker varies her production between flapping and the alveolar plosive.

Let’s now turn to the productions of the Brazilian Portuguese speaker learning English as a foreign language. Data here are from the subject RH, the same subject whose data we looked at when examining the production for “pet”.

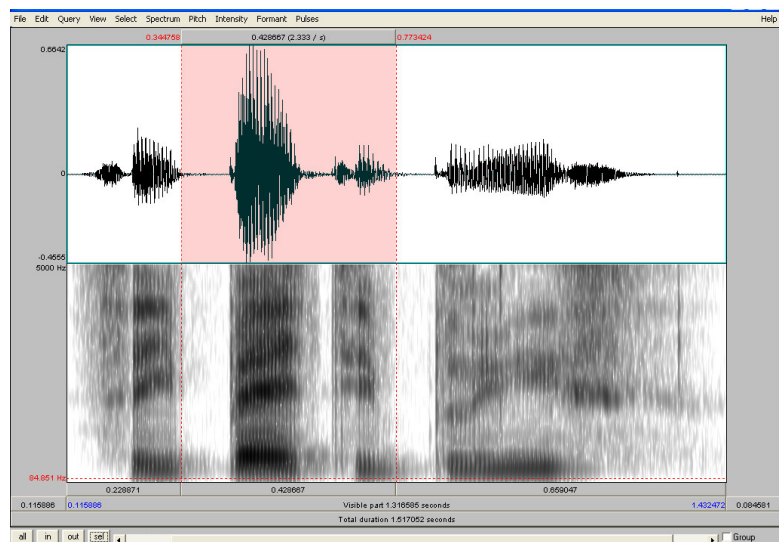


Figure 8a. The word "petty" produced by the Brazilian Portuguese speaker RH

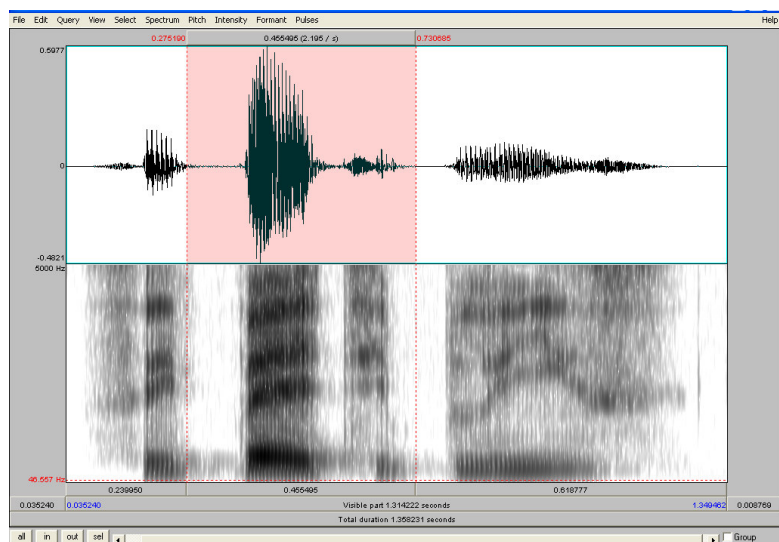


Figure 8b. The word "petty" produced by the Brazilian Portuguese speaker RH

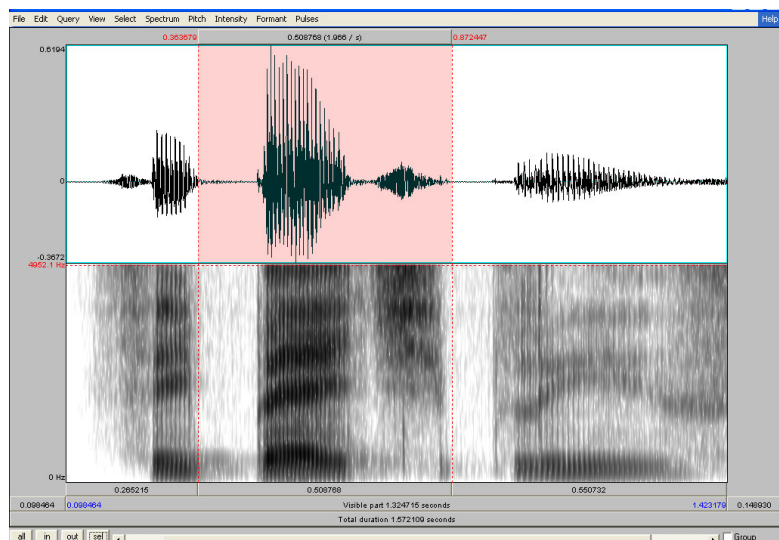


Figure 8c. The word "petty" produced by the Brazilian Portuguese speaker RH

The three productions for “petty”, spoken by the Brazilian Portuguese learner of English as a foreign language were shown above in order to observe that he also varies his production, but not in the same way the English native speaker GA did.

Figures 8a and 8b show that subject RH produced “petty” with an affricate consonant, instead of an alveolar plosive or a flap. This is not necessarily an error, since subject GA also produced [t] followed by frication noise, as we mentioned by purpose of Figure 7c. It is also worth observing that the duration of the frication noise is similar to both subjects⁷, i.e., the learner and the native speaker produced the same articulatory maneuvers. Once more, like in the case of the pair “luck/lucky”, the difference between the productions of the native speaker and the nonnative one seems to be prosodic, since the nonnative speaker produces the word “petty” as if it were a trochaic dissyllable.

Finally, Figure 8c exhibits a very interesting fact: the subject begins to produce the frication noise before producing the plosive [t], i.e., he starts making a big constriction in the vocal tract, then makes an obstruction and after that he releases the obstruction, producing a constriction again. We can say, then, that he anticipates the articulatory maneuver that results in the frication noise. It is reasonable to think that he is trying to adjust the articulatory gestures necessary to produce the sound sequence as in the native speech.

4. A dynamical approach to explain what learners do

In the light of the data examined in the last section, it must be clear now that: a) Brazilian Portuguese speakers learning English as a foreign language do not insert a segment nor replace one segment by another one in the cases of plosives in word-final coda observed here; b) learners do make a difference between the members of pairs such as “luck/lucky; “pet/petty”, what means that they are able to produce plosives word-finally as well as to perceive the morphophonological difference underlying pairs such these; c) the difference between the native speech and the nonnative one, in this case, seems to be in the prosodic

⁷ The time axis for the six figures is very close, varying from 1,29 s to 1,33 s, so that a visual inspection can be more accurate.

domain⁸, since Brazilian Portuguese speakers tend to produce words such as “lucky” and “petty” as trochaic dissyllables, that is the most usual stress pattern in Brazilian Portuguese.

These points are all evidences that learners do not make “errors”. They are trying to produce the sound sequences the closest to the foreign language patten they can. Sometimes they exaggerate in their attempt, as it was the case with subject RA when producing the glottal effect in the word “luck”.

If it is not the case of inserting or deleting a segment or even replacing one segment by another, generative phonology cannot account for the facts explored here, because phonological models in this framework intend to deal with phenomena like those. Moreover, the data investigated in this paper leads to the conclusion that the main characteristic of the nonnative speech is that we can find in it the same acoustic events we find in the native speech, but they are sometimes longer than the same events in the native speech or they begin in a different time, comparing the succession of events in the native and in the nonnative speech. It is not the purpose of generative phonology to account for these phenomena. The question that arises then is: what kind of model could provide an adequate approach to the facts observed here?

We argue that dynamical models of speech production, such as Articulatory Phonology (Browman & Goldstein 1986, 1989, 1990, 1992) can account for the facts mentioned in this paper.

Articulatory Phonology does not assume the feature or the segment as its theoretical prime. Instead, it takes the articulatory gesture to be its unit. The articulatory gesture is at the same time an abstract and a concrete unit, derived from a mathematical equation that models the movement of a simple dynamic system of the type mass-spring. Articulatory gestures are also claimed to be pre-linguistic units, in the sense that they can be found in the speech of a child before it had acquired its mother tongue. Language acquisition process is then a fact that requires the child to learn how to relate the articulatory gestures in its speech just the way adults do⁹. Articulatory Phonology takes eight tract variables to engender all articulatory gestures that occur in all languages. These variables are produced by six articulators: lips, tongue tip, tongue dorsum, jaw, velum and glottis.

As a result of being derived from a mathematical equation that models the movement of a simple dynamic system, the articulatory gestures are units of intrinsic timing, differently from segments or features. Because of that, Articulatory Phonology assumes that gestures superimpose each other over time and the superimposition can result in different phasing relations between the gestures. These different phasing relations, in turn, are responsible for the gradiency that lays beyond the speech chain and that result, e.g., in facts such the one pointed out in section 2.2, by purpose of the way the subject GA produced the word “petty”: remember that Figure 7a showed an acoustic event similar to a flap and to an alveolar plosive at the same time, i.e., an acoustic event in between two categories.

Differences in phasing relations among articulatory gestures can also explain the production for “petty” by subject RH illustrated in Figure 8c, when he started frication noise, then produced a plosive and after that concluded the frication noise. Yet, differences in the way native speakers phase the tongue dorsum gesture that leads to the frication noise relatively to the tongue tip gesture that leads to the closure gesture in the plosive [t] and the way nonnative speakers phase the same gestures can account for the facts we examined in section 2 above. The phasing of the articulatory gestures can also account for the differences we noticed between the native and the nonnative speaker in the production of the pair “luck/lucky”.

⁸ This paper does not aim at investigating the prosodic influence of L1 on L2, although we recognize this is a very important topic that deserves an accurate study.

⁹ Perception must play an important role in this task, but it is not clear yet how perception acts in the adjustment of articulatory gestures in a child’s speech relatively to the adult’s speech.

Assuming phasing relations between articulatory gestures act in these cases, we can account for facts in a way other theoretical frameworks cannot do.

5. Final remarks

The acoustic analysis conducted in Section 2 showed that, differently to what literature usually claims about Brazilian Portuguese speakers learning English, they do not necessarily insert an epenthetic vowel in words ending with a plosive consonant. On the contrary, they are able to produce words with a plosive in the final coda, like an English native speaker does.

The difference between a native speech and a nonnative one, as we pointed out before, lies in the duration of the acoustic events and in their sequence. This fact cannot be adequately handled by theories of extrinsic timing, such as the phonological models inspired on Chomsky and Halle's generative phonology. That's because such models assume phonological facts to be based on phenomena like insertions and deletions of segments or features. This is clearly not the case here. That's why we argue in favor of a dynamic approach to explain what learners are in fact doing when they exhibit productions such as the ones we observed in this paper. A dynamical model, such as Articulatory Phonology, can explain what is happening taking into account the relationship between the articulatory gestures involved in a speech chain.

The consequence of assuming a dynamic view on L2 phonological acquisition is that learners have to be taught how to relate the articulatory strategies they already produce following the timing of these articulatory strategies in L2. It's not the case, then, to teach the students what they already know and what they are already able to produce. This perspective requires the teacher to see his student in a different way, i.e., not as one who does not know how to produce sounds or sequences of sounds, but as one who is capable of producing a number of articulatory gestures that are organized differently from language to language. The students need then to learn how to coordinate the gestures as it is in the L2 they are acquiring. Departing from the perspective that the student knows many things and that the learning process consists of improving this knowledge seems to make much more sense than a perspective that assume students have to learn everything, departing almost from zero, as is the case of the acquisition of L2 phonological level especially in the cases that L2 is not similar to L1, i.e., when they do not pertain to the same linguistic family.

References

- Albano, E. C., Moreira, A., Silva, A. H. P., Aquino, P., & Kakinohana, R. (1999). Um conversor ortográfico-fônico e uma notação prosódica mínima para síntese de fala em língua portuguesa. In E. Scarpa (Ed.), *Estudos de prosódia* (pp. 85-109). Campinas: Editora da Unicamp.
- Browman, C., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.
- Browman, C., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology Yearbook*, 6, 201-251.
- Browman, C., & Goldstein, L. (1990). Tiers in articulatory phonology. In E. Kington & M. Beckman (Eds.), *Papers in laboratory phonology I – between the grammar and the physics of speech* (pp. 341-376).
- Browman, C., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155-180.
- Chomsky, N., & Halle, M. (1991 [1968]). *The sound pattern of English*. Cambridge, Massachusetts: The MIT Press, 3rd. printing.
- Eckman, F. R. (2004). From phonemic differences to constraint rankings: Research on second language phonology. *Studies in Second Language Acquisition*, 26, 513-549.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

- Ferreira, A. P. (2007). Pet or petty? Diferenças entre palavras CVC e CVCV do inglês por aprendizes brasileiros: Uma análise acústica. Unpublished Master's thesis, Universidade Federal do Paraná, Curitiba Brazil.
- Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, USA: York Press.
- Fox, R., & Terbeek, D. (1977). Dental flaps, vowel duration and the rule ordering in American English. *Journal of Phonetics*, 5, 27-34.
- Massini-Cagliari, G. (1992). *Acento e ritmo*. São Paulo: Editora Contexto.