[ʃ] – [s] Accommodation in European Portuguese: an acoustic and perceptual study

Amália Andrade
Centro de Linguística da Universidade de Lisboa , Lisbon, Portugal
E-mail: a.andrade@clul.ul.pt

ABSTRACT

The present paper presents the results of an acoustic and perceptual study of [ʃ]-[s] accommodation in a European Portuguese (EP) dialect, based on intra-word material. It shows that stress plays an important conditioning role and offers some evidence in support of the idea that perceptual effects may contribute towards a reinforcement of the ongoing change.

1 INTRODUCTION.

In the Lisbon dialect of European Portuguese (EP), sequences constituted by a syllable-final palato-alveolar /S/ and an onset (denti)-alveolar /s/ may "simplify" as a singleton palato-alveolar fricative, both intra-word and across word boundary. The phenomenon raises interesting questions since it violates canonical coda-onset relation and does not support directionality predictions [1], which other EP dialects and languages seem to support [2]. Its expansion seems to be relatively recent, considering that about a century ago coda fricatives were weak, in the Lisbon dialect [3].

The phonetic study of casual speech phenomena offers relevant information for a better understanding of phonological change. It is commonly accepted that an important source of change is found, in many cases, in speech production. Listeners' misperception of the variable phonetic outputs is also proposed as a possible source of sound change. The present study aims at establishing some of the conditions that may favour the emergence of [ʃ] and at looking for evidence in support of the idea that perception may contribute to sound change and its spread.

2 ACOUSTIC EXPERIMENT

2.1 METHOD

Our speech material was constituted by realizations of underlying sequences of coronal fricatives occurring within a word. The material in question was extracted from 6 relatively short declarative sentences read by 5 Lisbon speakers, one man (S1) and four women (S2-S5) of ages ranging between 23 and 35. These sentences are part of a corpus of 223 sentences, which were repeated four times each, in random order. The subjects were asked to read in a natural manner and at their normal rate. Recordings were made in an anechoic chamber.

The material utilized pertains to 3 forms of the verb "nascer" ('to be born'): (a) two in which /s/ is the onset of the stressed syllable, "Nasceram" ([nɐ]/S´s/[erɐ̃w̃]) and "nasceu" ([nɐ]/S´s/[ew]), and (b) one in which coda /S/ belongs to the stressed syllable, "nascem" ([´na]/Ss/[ɐ̃j̃]). Quality of the preceding and following vowels (V_1 and V_2) was not kept constant partly due to the rising process EP unstressed vowels undergo. Position in the sentence was varied: "Nasceram" is sentence initial while "nasceu" and "nascem" occur medially, in a syntactically unmarked order.

A spectral and temporal analysis was carried out. With respect to the latter, fricative durations as well as total utterance durations were measured. The latter were utilized to calculate average articulatory rate values (number of syllables/utterance duration). Fricative signal subsegmentation was based on the detection of (quasi-) stationary regions, using spectrograms and dft spectra (with a 256pt Hanning window; sampling rate=16kHz). Spectral analysis involved measurement of the lower cutoff frequency and dominant spectral peak frequencies and amplitudes by means of 15ms-averaged dft spectra (256pt window) placed at the beginning (starting 10ms from fricative onset) and the end (ending 10 ms from fricative end) of the fricative regions. A finer spectral analysis was carried out, namely when there was evidence of anticipatory rounding (as in some instances of the "nasceu" set) and/or if there was a significant indication of blending, namely when the initial and final dominant peak locations pointed to a single, but non-[ʃ] and non-[s] fricative (the latter are not to be expected in this dialect). One of the stages of the analysis was precisely a classification of the fricative tokens as [ʃ], [ʃs] or ambiguous (A). For this purpose, we also analysed singleton /ʃ/s and /s/s occurring at word onset, and F4 at midpoint of V_1, for all speakers. Fig. 3 includes dfts of A-type items.

2.2 RESULTS

According to the present analysis, our 5 speakers' /Ss/ realizations involve 46% [ʃ]-, 13% A- and 41% [s]. 59% of

 ISBN 1-876346-48-5

the total [ʃ] set corresponds to "nascem", 26% to "nasceu" and 15% to "Nasceram". Relative to the number of items per condition (20), 80% of the "nascem" items are [ʃ], whereas 80% of the "Nasceram" tokens are [ʃ]s; for "nasceu" we have found 65%[ʃ]s.

The number of [ʃ]-outputs also varies across subjects, with S1 and S4 representing its maximal and minimal values, respectively (cf. Fig.1).

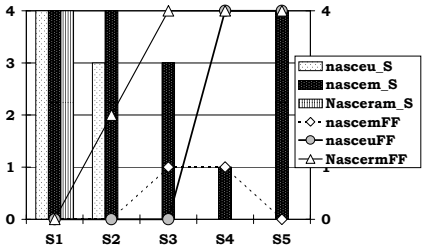

Figure 1: Acoustic results: number of intra-word [ʃ]s (columns) and [ʃs]s (icons) per subject (S1-S5 from left to right).

Mid-utterance position together with "stressed coda" ("nascem") is the condition, which favours [ʃ] outputs for all subjects, though in a varying degree. [ʃ] occurrence in other conditions is restricted to S1 and S2.

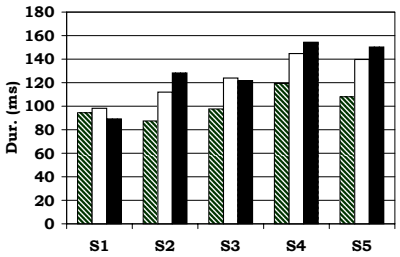

Figure 2: Average durations for /Ss/ in "nascem" (barred), "nasceu" and "Nasceram" (black) for S1-S5 (from left to right).

/Ss/ average durations of "nascem" are the shortest of the three conditions for all subjects except S1, whose duration differences are not significant (Fig. 2). Subsegmentation of the fricative regions shows that the durations of coda-[ʃ] regions of phonetic [ʃs]-type sequences range between less than 30ms and more than 70ms, depending on the speaker (they are shortest for S3 and longest for S5) but they are always shorter than the [ʃ] realizations of "nascem".

Leaving S1 aside, the duration differences associated with stress position ("nascem"/"nasceu") are statistically significant (Mann-Whitney test) for all subjects. Implementation of stress [5], that is of the prominence relation between a stressed and a following unstressed syllable, in turn, seems to play a major role, both in terms of its generality and in terms of the relative weight it has for most speakers. On the other hand, influence of utterance position ("nasceu"/ "Nasceram") is speaker dependent: S2 is the only speaker for whom this factor is associated with significantly different duration values.

Average articulatory rate indexes (8.6 for S1 vs 7.4, 7.4, 6.9 and 6.8 for and S2, S3, S4 and S5, respectively) and total average durations (94 ms for S1, vs 109, 115, 140 and 133 for S2, S3, S4 and S5, respectively) indicate that the highest frequency of occurrence values correspond to the fastest speaker and the lowest frequency of occurrence values correspond to one of the 2 slowest speakers (S4). It is possible that rate partly accounts for S1's attaining 100% [ʃ] outputs; maybe it has something to do with S2's [ʃ] outputs, but it cannot offer an explanation for S5's 100% [ʃ]s. The present results, however, also allow for the hypothesis that in the case of S1 and S5, coarticulation of the tongue dorsum and tongue front for the coda and onset fricatives [1] is governed categorically by higher level prosodic factors; the process is of a gradual nature for others. With respect to S1, we may be more radical and hypothesize that /ʃ/ substitution for /Ss/ has taken place in his lexical representation of the verb "nascer"; rate, then, would not be a primary factor.

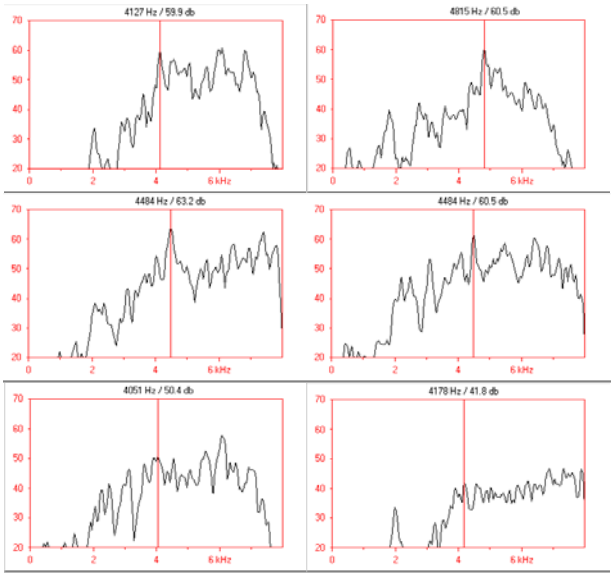

Figure 3: Time averaged dft spectra (at end of fricative) of (i) S2 – onset-[ʃ] (top left), (ii) A-type: S4 - "nascem" item (top right), S2 – 2 "Nasceram" (middle) and 1 "nasceu" items (bottom left), S3 – 1 "nasceu" item (bottom right). Cursor between 4 and 5 kHz.

Evidence of gradience in the acoustic manifestations of /Ss/ sequences has been found for S2 in the "Nasceram" context, S4 in "nascem", and S3 in "nasceu" (Fig.3), which is suggestive of articulatory blending. One possible explanation for these speaker's "nascem" outputs is that articulatory overlap and front articulation reduction takes place. Durations and spectral data corresponding to these speakers' "Nasceram" realizations, together with the F4 values of V_1 and the spectral data from singleton onset [ʃ]s and [s]s indicates that tongue fronting weighs more in S3's productions than in those of the other speakers and that S4 aspirates significantly more than any other speaker.

The results featured in Fig.1 and the spectral characteristics of S2's, S3's and S4's /Ss/ realizations of "nasceu" (rounded context; Fig. 3) show that rounding coarticulatory effects are highly speaker dependent.

ISBN 1-876346-48-5

3 PERCEPTION EXPERIMENT

3.1 METHOD

Two identification tests were prepared with stimuli excerpted from the material previously analysed. The stimuli of the two tests differed basically with respect to the inclusion (T1) or the exclusion (T2) of vocalic context preceding and following the fricative segments.

The durations of the vocalic sections of these stimuli were defined roughly in proportion with the actual durations of the source contextual vowels, in most cases. In two of S3's realizations of "nascem", in the absence of V_2 manifestation, a segment of the following nasal consonant was also included. 3 [ʃ]-type items and 3 [ʃs]-type items from 3 different speakers, included a 20ms portion of the C-V_1 transition. It was thought that the presence of weak cues to lexical forms might reinforce the distinction between the two experimental conditions: one more low-level based and the other more prone to involve higher order processing. Transients at stimuli onsets and offsets were avoided. Each test included 600 items (60 stimuli x10). These were ordered randomly.

25 subjects were used for the present study. Most listeners were Lisboners, others were also speakers of dialects with the EP standard fricative system. Subjects were asked to identify the stimuli as [ʃ] or [ʃs]. An interview followed the test sessions, but we will not refer to this type of information in the present paper. Some subjects repeated the tests a second time, for assessment of consistency in the responses and possible learning effects.

3.2 RESULTS

Comparison of the average rates of [ʃ] identification and the results of our classification of the acoustic data as [ʃ]s (vs non-[ʃ]) (Fig. 4) shows great proximity between the latter and the results of T2. The middle right and bottom left dfts of Fig. 3 correspond to the stimuli that are [ʃ] in T2 (average results) but non-[ʃ] in our analysis. It was thus adequate to assign T2 a (gross) reference status in the interpretation of the results obtained in T1.

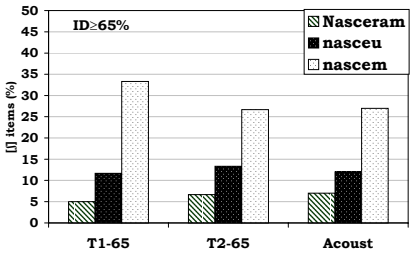

Figure 4: "Nasceram" (barred), "nasceu" (white spots), "nascem" (black spots): percent rates of [ʃ] detection obtained (i) in the acoustic analysis (right) and (ii) in Tst1 (left) and Tst2 (middle). Averages of 25 subj.s (id. ≥65%).

When average identification responses are plotted as a function of the frequency of the lower dominant peak at the end of the fricative (Prome), the [ʃ]:[ʃs] identification boundary lies between 4000 kHz and 4500 kHz (Fig. 5), which is in accordance with previous findings for EP [6,7] and other languages. The frequency span of the boundary region along Prome is clearly conditioned by contextual characteristics. However, influence from segmental contexts per se cannot account for the differences found between the two tests.

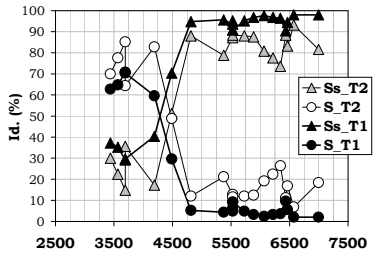

Figure 5: [ʃ] and [ʃs] % identification scores (av.s for 25 subj.s) corresponding to the "Nasceram" stimuli, in T1 (black) and T2. Prome frequencies (2.5-7.5kHz) on *x* axis.

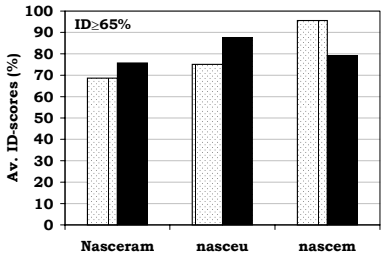

Figure 6: Average [ʃ] identification scores (%) for 25 subjects (id-criterion ≥65%) in T1 (spotted) and T2 (black). Pairs of columns, from left to right: "Nasceram", "nasceu" and "nascem".

The total average identification rates corresponding to "Nasceram", "nasceu" and "nascem" reveal that more "nascem" stimuli were perceived as [ʃ]s in T1 than in T2, and that the opposite applied to the other stimuli (Fig. 4). A similar trend is found in the average identification scores (identification criterion ≥65%): higher average [ʃ] scores were attained for the "nascem" stimuli in T1 than in T2 and the opposite was found for the "Nasceram" and "nasceu" stimuli (Fig. 6).

Examination of the individual results reveals important facts that are masked by the averaged results. Here, we will lower our identification threshold to ≥60%. To start with we find that 40% of the listeners, group1, identify all "nascem" stimuli as [ʃ] in T1 and do not apply this "generalization" mode to the other contexts. This in itself accounts for what we have just observed, to a great extent. Another 16% of the subjects, group2, identify all "nascem" stimuli as [ʃ] in T1 *and* in T2. Finally, we have group3, constituted by listeners who also "generalize" in one or both of the other contexts: 16% of the remaining 44% of the subjects identify all "Nasceram" and "nasceu" stimuli as [ʃs], and another 12% apply this categorization to "Nasceram" (a total of 28% classify all "Nasceram" items as [ʃs]); a few other subjects (8%) approximate the latter subgroup. Fig. 7

ISBN 1-876346-48-5

contains response curves for "Nasceram" in T1 and T2 of representatives of group1 (top, Subj-A) and of group3 (middle, Subj-C). The bottom curves, correspond to "nasceu" of Subj-C, and show that an effect of "learning" may take place from a session of T1 to another occurring one week later. We should clarify that the occurrence of a learning effect in T1 is not specific of group3 subjects.

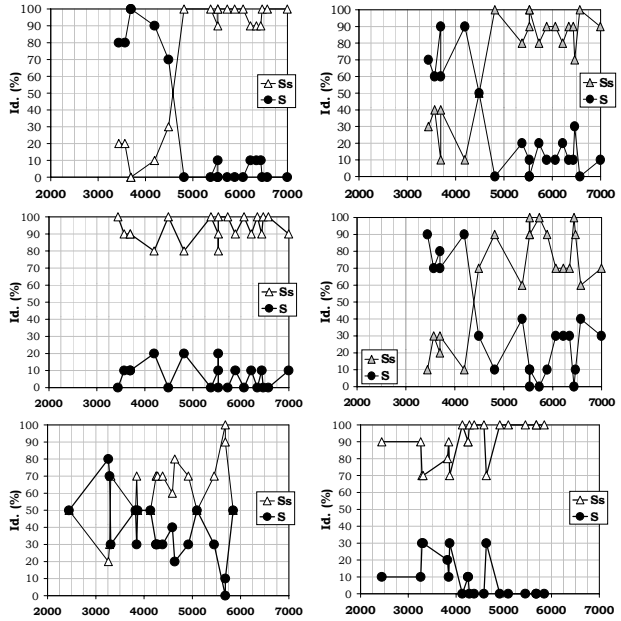

Figure 7: [ʃ] (black circles) and [ʃs] (triangles) identification curves for (i) "Nasceram" stimuli - T1 (left) and T2 (right), for 2 subj.s (A –top and C-middle); (ii) subj. C's "nasceu" stimuli identific. curves (bottom) in two sessions of T1 (1st – left; 1 week later – right). *x* axis: Prome (Hz).

The present data do not allow us to be very assertive with respect to what may underlie these three types of perceptual behaviour, but it is clear that groups 1 and 2 rely more on the phonetic properties of the fricative stimuli than group 3, in T1. It is conceivable that secondary properties of the stimuli in question, such as their relative shortness plus aspiration (in the case of S4), and/or stronger blending (cf. top right dft in Fig. 3) compensation of higher spectral peaks by rounding effects (Fig. 3, bottom), play a part. In the case of group1, the combined effect of frequency distribution is to be considered (cf. initial paragraph of 2.1). Group2 may also be thought to have broader phonetic category boundaries for [ʃ] than other listeners. It is possible, too, that an identification strategy based on the presence/absence of an [s] results in the greater acceptability of stimuli as [ʃ] for groups1 and 2. The latter factor, however, is more directly associated with the experimental design itself and would therefore require better control. Unlike groups 1 and 2, who seem to focus on low-level properties, group 3 seems to operate at a higher level, combining use of language knowledge (e.g. stress) and contextual information. In this "mode", the relative importance of frequency distribution would be significantly reinforced. In fact there seems to be a strong relation between the pattern of responses of this group and the pattern of frequency distributions of [ʃ]s and [ʃs]s across contexts (80% "nascem" [ʃ]s, 80% "Nasceram" [ʃs]s and 65% "nasceu" [ʃs]s). Thus, different expectancy effects resulting from differences in perceptual strategies would partly account for group1's and group3's patterns of responses.

CONCLUSION

This study shows that occurrence of [ʃ]-outputs in the realization of /Ss/ in EP is favoured when the coda fricative belongs to the stressed syllable of a word rather than to an unstressed one. It is case, then, where higher-level prominence relations invert the expected directionality of [ʃ]-[s] coarticulation [1] and override the canonical coda-onset relation. Speakers display different patterns and it is conceivable that differences in lexical representation be involved.

The perceptual results reflect a tendency from the part of listeners to "misperceive" non-[ʃ]s as [ʃ] more, when the probability of occurrence of the latter is highest, that is when the coda fricative belongs to the stressed syllable. The results are compatible with the idea that the phenomenon in question may be reinforced perceptually and that higher-level (including linguistic knowledge) and low-level factors contribute towards phonetic change and its spread.

ACKNOWLEDGEMENTS

The author is most grateful to H. Onofre for his technical support with the recordings, to R. Berkovitz, author of the software used in the perceptual experiments, as well as to the 30 subjects who so patiently participated in the experiments. Access to the recording facilities of the CAPS (Technical U. of Lisbon) is also thankfully acknowledged.

REFERENCES

[1] D. Recasens and D. Pallarés, "Coarticulation, assimilation and blending in Catalan consonant clusters", *J. of Phonetics*, vol. 29, 3, pp. 273-302, 2001.

[2] F. Nolan, T. Holst and B. Kühnert, "Modelling [ʃ] to [s] accommodation in English, *J. of Phonetics*, vol. 24, 3, pp. 113-137,1996.

[3] A. Gonçalves Viana, "Essai de phonétique et de phonologie de la langue portugaise d'après le dialecte actuel de Lisbonne", *Romania*, *12*, pp.29-98, 1883; also in A. R. Gonçalves Viana, *Estudos de fonética portuguesa*, Lisbon: Imprensa Nacional, Casa da Moeda, pp. 83-152, 1973.

[4] M. R. Delgado-Martins, *Sept études sur la perception*, Linguística 10, Lisbon: INIC, 1986.

[5] J. Ohala, The phonetics of sound change. In C. Jones, ed., *Historical Linguistics: Problems and Perspectives*, London: Longman-Academic, pp.237-278, 1993.

[6] F. Lacerda, "Acoustic perceptual study of the Portuguese voiceless fricatives", *J. of Phonetics*, vol. 10, pp. 10-22, 1982.

[7] L. Jesus, *Acoustic properties of European Portuguese Fricative consonants*. PhD Diss., U. Southampton, 2001.