PERCEPTION AND PRODUCTION OF ITALIAN PLOSIVES BY AUSTRIAN LEARNERS

H. Grassegger

Abteilung für Phonetik
Institut für Sprachwissenschaft, Graz, Austria

ABSTRACT

Two experiments on the differentiation of bilabial plosives were carried out to test the hypothesis that articulatory difficulties have a perceptual correlate which may function as a diagnostic tool. An identification test on a synthesized 10-stimulus continuum with different VOT-values ranging from -80 to +64 msec showed listener specific subcategorisations. These subcategorisations are reflected in the individual voicing and aspiration parameters of the subjects' realisations of bilabial plosives in initial prevocalic position.

0. INTRODUCTION

0.1. Previous studies on the production of plosives by speakers having an Austrian variety of German as their native language (see [1], pp. 209 - 234) yielded the following facts: a) Lenis plosives (/b,d,g/) in initial prevocalic position were produced without voicing. b) Articulatory differentiation of initial prevocalic lenis and fortis plosives resulted in shorter voicing lags for lenes, longer voicing lags for fortes. Average differences amounted to 10 msec for labials (/b/: 20 msec - /p/: 30 msec), 20 msec for dentals (/d/: 25 msec -/t/: 45 msec), and 55 msec for velars (/g/: 30 msec - /k/: 85 msec).

0.2. It is the differentiation of (voiced) lenis and (unaspirated) fortis plosives which is specifically problematic for Austrian (as well as for German) learners of Italian. Measurements of Italian voiced plosives (here again: in initial prevocalic position) revealed a considerable voicing lead around 80 msec (see [1], p. 149).

1. HYPOTHESIS

1.1. It is now hypothesized that the traditional difficulties which Austrian speakers show in adequately pronouncing Italian voiced and voiceless plosives of the same place of articulation are not only to be explained by differences in the sound pattern of source and target language but that these difficulties have a perceptual correlate. That is to say that an individual learner's perceptual habits would presumably reveal his efficiency in articulation.

1.2. In this study perception and production of the contrast between (initial prevocalic) /p/ and /b/ by Austrian subjects learning

TABLE 1: Perception test

Stimulus nr.	1	2	3	4	5	6	7	8	9	10
VOT (msec)	-80	-64	-48	-32	-16	0	+16	+32	+48	+64
					GERMAN TEST GROUP					
Listener 1	48	49	47	46	42	39	*30	*19	7	4
Listener 2	50	50	40	*31	*24	18	15	2	0	0
Listener 3	47	49	47	41	40	38	32	/ 18	16	2
Listener 4	48	44	46	45	39	40	37	/ 16	5	0
Listener 5	46	47	45	43	36	34	*31	*30	10	0
Listener 6	49	48	46	44	41	38	*28	18	3	2
Listener 7	49	50	49	32	*20	16	10	2	0	0
Listener 8	48	47	42	41	42	40	*30	16	0	0
Listener 9	50	50	48	*27	*24	14	13	0	1	0
Listener 10	47	44	44	43	38	34	33	/ 11	2	0
					ITALIAN CONTROL GROUP					
Listener 1	50	50	38	*19	12	5	3	0	0	0
Listener 2	49	50	34	/ 16	8	2	4	1	0	0
Listener 3	50	50	35	*22	10	4	2	0	1	0
Listener 4	49	49	32	*20	15	7	3	0	0	0
Listener 5	50	49	37	/ 17	11	3	4	0	0	0

Italian is examined. The acoustic dimension that has been decided to be most useful for this purpose is voice onset time (VOT).

2. PERCEPTION TEST

2.1. Test Structure. A series of 10 /Ca/ stimuli with VOT values ranging from -80 to +64 msec in 16 msec increments (Fo 125 Hz) were synthesized using the synthesis system Flexivox (for details on the system see [2]). Spectrum and intensity of the release burst (320 msec prior to the end of the vowel) represented an average of values used in the common text-to-speech synthesis of bilabial (/b-/ and /p-/) stops. Each synthesized stimulus was recorded on a Revox tape recorder (B 77MK II) 5 times in randomized order with a fixed interstimulus interval of 2 seconds. The resulting block of 50 stimuli was then copied 10 times with a 10-second pause after each block (thus yielding 50 tokens of each stimulus) and administered to ten Austrian learners of Italian, whose experience of Italian ranged from one to four years (German test group) and to five native listeners from Northern Italy (Italian control group). The stimuli were presented binaurally over headphones (PIONEER SE-305) at about 70 dB SPL(A). A practice block consisting of two repetitions of the 10 stimuli in ascending order of VOT was presented before the experiment for familiarization with the synthetic voice. The subjects were told that each stimulus represented the beginning of either Italian 'palla' ("ball") or 'balla'("bale"). They had to indicate their judgement by circling 'palla' or 'balla' on an answer sheet. Moreover, they were asked to guess in case of uncertainty.

2.2. Test Results. The results for both groups of listeners are given in table 1 as the number of /b/-identifications for each stimulus. Values not significant according to a chi-square test ($19 \le x \le 31$, $n = 50$, $p = 0.05$) are marked by an asterisk. In case of lacking asterisked values a slash separates significant /b/-judgements (left of the slash) from significant /p/-judgements (right of the slash). In all other cases numbers preceding asterisked values represent significant /b/-identifications, those following significant /p/-identifications. Diagrams of the individual results rendered the basis for interpolation of VOT-values that represent the 50% crossover of judgement. With all Italian listeners and with three of the German listeners these crossover VOT-values above which more than 50% of the tokens of one stimulus were judged to be voiced lie within the voicing lead, for the remaining German subjects it reaches far into the voicing lag (cf. table 2).

TABLE 2: Interpolated cross over VOT-values

German test group				
L 1	L 2	L 3	L 4	L 5
+22	-22	+22	+21	+36
L 6	L 7	L 8	L 9	L10
+25	-23	+23	-19	+24
Italian control group				
L 1	L 2	L 3	L 4	L 5
-37	-40	-36	-39	-38

It was for this reason that the German test group was finally divided in two subgroups according to their cross over points: subgroup 1 (comprising listeners 1, 3-6, 8, 10) with crossover in the voicing lag, subgroup 2 (listeners 2,7,9) with cross over in the voicing lead. In figure 1 the mean values of /b/-judgments for the two German subgroups and the Italian control group are plotted for comparison.

The subcategorisation of the 10-stimulus continuum is seen to be shifted to the right by German listeners compared to the Italian group, more strongly so for subgroup 1 than for subgroup 2. As a result, stimuli with no voicing lead or even with a voicing lag up to about 24 msec are perceived as acceptable realisations of the beginning of Italian 'balla' by subgroup 1. Subgroup 2 more rigorously rejects stimuli without any voicing lead as 'ba' but does not reach the cross over values of the Italian listeners.

3. PRODUCTION TEST

For the production test the same ten subject of the German test group read a list of ten Italian words beginning with voiced and voiceless bilabial stops which included the two words 'balla' and 'palla' and repeated it five times. The subjects' speech was recorded on a tape recorder (Revox B 77MK II). The five realisations of 'balla' and 'palla' respectively were then analyzed by means of a digital oscilloscope (Nicolet 3091), which allowed easy measurement of VOT-values in steps of 0.2 msec. Measurement reliability was assessed by remeasuring ten randomly selected realisations. The mean difference between the two measurements of each stop was 1.2 msec with a range of 0.8 msec.

The results of the production test are given in table 3 as mean values of the five /b/- and /p/-realisations.

Only those speakers who are members of the German subgroup 2 according to the perception test show considerable voicing leads in their /b/-realisations.

TABLE 3: Production test VOT-values in msec

	/b/		/p/	
Sp	avg	sd	avg	sd
1	+20.6	2.4	+30.8	2.4
2	-27.8	3.2	+27.8	1.9
3	+21.4	2.1	+27.2	1.5
4	+19.6	3.8	+24.2	1.6
5	+26.6	3.2	+38.8	1.6
6	+18.8	2.4	+32.8	1.9
7	-37.6	3.9	+26.4	1.4
8	+23.4	2.0	+30.8	2.4
9	-32.4	2.4	+19.8	1.9
10	+22.4	2.9	+37.6	2.7

4. CONCLUSION

The results of the perception and production test on the one hand show the expected difference in categorisation ability between native speakers and learners of the language and, moreover, clear differences among the learners themselves. On the other hand they demonstrate most convincingly that well-established perceptual categories are more likely to be accompanied by more acceptable production.

Thus, our hypothesis on the interrelationship between perceptual and articulatory abilities seems to be proved. The link between perceptual categories and articulatory differentiation is an encouraging indication that suitable perception tests provide a powerful diagnostic tool in pronunciation teaching.

5. REFERENCES

[1] GRASSEGGER, H. (1988), *"Signalphonetische Untersuchungen zur Differenzierung italienischer Plosive durch österreichische Sprecher"*, Hamburg: Buske Verlag.

[2] OLASZY, G., GORDOS, G. (1988), "Die Anwendungen des Flex-Deutsch Sprachsynthese-systems in phonetischen Forschungen", *Hungarian Papers in Phonetics, 19*, 34 - 46.

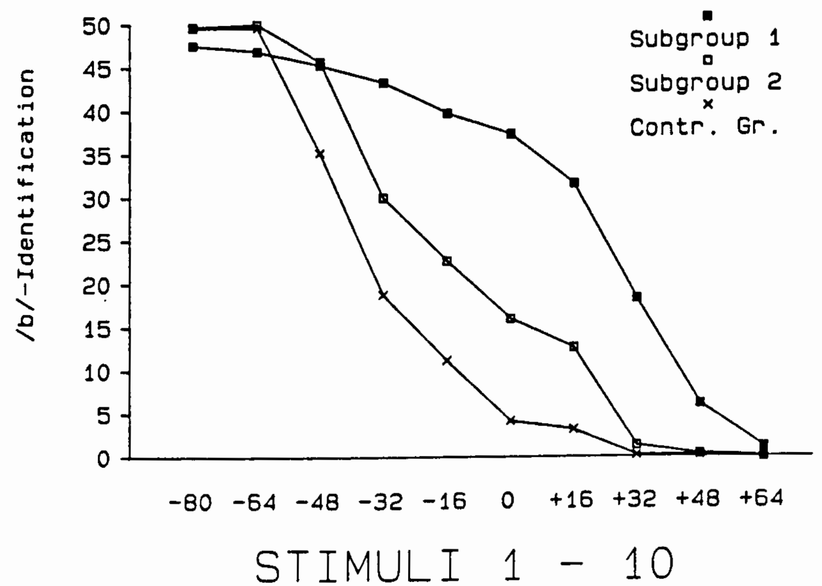

FIGURE 1: Mean values of /b/-judgements