Place cues for nasal consonants with special reference to Catalan

Daniel Recasens

Department of Linguistics, University of Connecticut, Storrs, Connecticut 06268 and Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06510

(Received 15 October 1982; accepted for publication 20 December 1982)

This study investigates the perceptual contributions of formant transitions and nasal murmurs to the identification of the unreleased Catalan nasal consonants [n], [ɲ], [ŋ] (alveolar, palatal, velar, respectively) after [a] in absolute final position. Transition and murmur patterns were synthesized and varied simultaneously and systematically by interpolating between optimal values obtained from spectrographic analysis of natural speech. Catalan subjects were asked to identify the synthetic stimuli as [n], [ɲ], and [ŋ]. The main findings were: (1) Although transitions provided more effective cues for place of articulation than murmurs, the murmurs did make a significant contribution to the [ŋ]–[n] distinction. (2) The cue value of the transitions ([ɲ] > [n], [ŋ]) was inversely related to that of the murmurs ([ŋ], [n] > [ɲ]). It is concluded that static and dynamic place cues for nasals in an [aC#] context are perceptually integrated with reference to the typical pattern of production of these consonants.

PACS numbers: 43.70.Dn, 43.70.Ve

INTRODUCTION

It is widely accepted that certain spectral characteristics of nasal murmurs mark nasal consonants as a class, thus acting as *manner cues*. Phoneticians agree on the relevance of the following distinctive spectral traits of nasal murmurs (Delattre, 1958; Fujimura, 1962): a low first formant ($N1$) at 250–300 Hz, with higher intensity than the upper formants; an antiformant (NZ), varying in frequency with place of articulation; a set of weak formants ($N2$, $N3$, $N4$,...) at 300–4000 Hz, with large bandwidth (bw) values; an overall lower intensity level than vowels. Other less important manner cues are: nasalization effects upon the neighboring vowels; nasalized releases; $F1$ transitions which are less pronounced than for non-nasal stops.

While these manner cues are well known, the evidence concerning *place cues* for nasals is less complete. Although the role of formant transitions has been studied (Liberman *et al.*, 1954; Nakata, 1959; Hecker, 1962; Garcia, 1966, 1967a, b; Miller and Eimas, 1977; Larkey *et al.*, 1978) very little attention has been paid to nasal murmurs as place cues in perception experiments. For instance, one might expect transitions and murmurs to complement each other in the perception of nasals in absolute final position, given the fact that transitions alone are poor cues in that environment (Liberman *et al.*, 1954). This prediction is examined in the present study with respect to final nasal consonants preceded by [a].[1] A perceptual interdependence of transitions and murmurs would suggest that place cues for nasals, at least in absolute final position, are perceived in an integrated fashion. Thus the process of place perception for nasal stops would not differ essentially from that for non-nasal stops (Dorman *et al.*, 1977) or fricatives (Whalen, 1981), in the sense that stationary and dynamic cues contribute simultaneously to place identification.

Acoustic analyses and perceptual studies of place cues for nasal consonants have considered released or unreleased nasal murmurs and formant transitions separately as well as in combination. A brief summary of findings relevant to the present study follows.

A. Released/unreleased nasal murmurs

Analysis data on nasal murmurs corresponding to [m], [n], [ɲ], [ŋ] in several languages—Czech, German (Romportl, 1973), English (Fujimura, 1962), Hungarian (Magdics, 1969), Polish (Dukiewicz, 1967), Russian (Fant, 1960), and Swedish (Fant, 1973)—reveal systematic differences in spectral parameters that might be perceptually relevant (> = higher frequency than):

$N1$ frequency values: [ŋ] > [ɲ] > [n] > [m];

$N1$ bandwidth values: [ŋ] > [ɲ],[n],[m];

NZ frequency values: [ŋ] > [ɲ] > [n] > [m].

The differences in $N1$ values are presumably related to differences in the size of the coupling section at the velopharyngeal passage and of the pharyngonasal tract. Contrastive NZ (nasal antiresonance) values are consistent with cross-category differences in oral cavity size behind the tongue constriction. Moreover, NZ falls close to a specific NF (nasal formant) that varies according to place category as follows: $N2$ for [m], $N3$ for [n], $N4$ for [ɲ], and $N4$ or higher NF for [ŋ].

Information about the perceptual relevance of murmurs as place cues is provided by experiments in which murmurs were presented for identification either in isolation (natural speech: Malécot, 1956; Dukiewicz, 1967; synthetic speech: House, 1957; Nakata, 1959) or directly attached to a steady-state vocalic portion (typically [a] or [æ]) without formant transitions (natural speech: Malécot, 1956; Henderson, 1978). Released (Malécot, 1956; Henderson, 1978) as well as unreleased (Henderson, 1978) murmur portions were tested. The results show the following trends:

[m] murmurs (English) were quite accurately identified regardless of whether they were isolated or vowel-attached and released or unreleased. Experiments with natural speech

gave better identification scores (80%–100%) than experiments with synthetic speech (65%–85%).

[n] murmurs (English) gave 50%–60% correct responses in all conditions. There was an important exception: naturally spoken released murmurs in Henderson's experiment (97% correct).

[ɲ] murmurs (Polish) gave no [ɲ] judgments but were mainly identified as [ŋ] (Dukiewicz, 1967).

[ŋ] identification results (English) were poor in the isolated condition (House: 62%; Nakata: 41%; Malécot: 12%) as well as after a steady-state vowel (Malécot: 34%). However, Henderson found a higher percentage of correct responses (75%–90%) using natural speech stimuli, independent of release.

The effectiveness of released or unreleased [m] murmur in place identification may be related to the particular low $N1$ and NZ frequency values within the overall low murmur spectrum (see also Delattre, 1958). The perceptual distinction between [ŋ] versus [n], [ɲ] is probably cued by high versus low $N1$ and absence versus presence of NZ in the central region of the murmur spectrum (see also Ohala, 1975).

B. Formant transitions

Formant transitions can show frequency decrease (falling or negative), frequency increase (rising or positive), or no frequency variation (steady) from vocalic steady state to transition onset/offset. (Note that an initial transition rising in frequency is classified as "falling" by this definition.) Analysis data on formant transitions for [m], [n], [ɲ], [ŋ] following or preceding [a] in several languages—Hungarian (Magdics, 1969), Italian (Vagges *et al.*, 1978), Polish (Dukiewicz, 1967), Russian (Fant, 1960)—show several relevant characteristics:

$F1$ transitions are consistently falling, more for [ɲ] than for [m], [n], and even less for [ŋ].

$F2$ transitions are consistently rising (by 500–1000 Hz) for [ɲ]; rising or flat for [n]; slightly rising, falling, or flat for [ŋ]; consistently falling (by 50–500 Hz) for [m].

$F3$ transitions are generally falling for [m], [n], [ŋ], and rising for [ɲ].

With respect to $F1$ values, the cross-category differences are mainly due to the relative increase in back cavity size of the vocalic relative to the consonantal configuration. While $F2$–$F3$ transitions for [m] are due to a complete labial constriction, those for the other categories reflect the change in front cavity size from [a] to the nasal consonant. Moreover, lower $F1$ and higher $F2$–$F3$ transition endpoint values for palatals than alveolars are presumably related to a greater narrowing of the palatal channel.

Information about the perceptual relevance of transitions as place cues is provided by experiments in which transitions were attached to a steady-state vowel (typically [a] or [æ]) without murmur portion (natural speech, VC structures: Henderson, 1978; Dukiewicz, 1967) or with a fixed nasal murmur (synthetic speech, CV or VC structures: Liberman *et al.*, 1954; Nakata, 1959; Hecker, 1962; Garcia, 1966, 1967a,b; Miller and Eimas, 1977; Larkey *et al.*, 1978). With respect to natural VC speech stimuli, while Polish [ɲ] identification is nearly perfect given a set of appropriate formant transitions (Dukiewicz, 1967), the number of correct responses is lower for English [an] (80%), [am] (60%), and very low for [aŋ] (15%–30%) (Henderson, 1978). Results obtained with synthetic VC stimuli show more satisfactory identification of [m], [n] than [ŋ] (Liberman *et al.*, 1954) or similar percentages of correct responses for all three categories (75%–80%, Larkey *et al.*, 1978).

The absence of [ɲ] in English explains the following findings in synthetic speech experiments with English listeners: $F2$ transition frequency ranges categorized as [ŋ] (+120 to +1200 Hz) correspond, to a great extent, to analysis values for $F2$ of [ɲ] in other languages (+500 to +1000 Hz) whereas analysis values corresponding to $F2$ for [ŋ] coincide, to a large extent, with values categorized as [m] (−500 to +800 Hz) or [n] (−500 to +360 Hz). On these grounds, it seems that stimuli with [ɲ]-like transitions are interpreted as [ŋ] by English listeners (although they may be categorized as [ɲ] by speakers of other languages) while stimuli with [ŋ]-like transitions, in the absence of an appropriate negative $F3$ and some [ŋ]-like murmur spectrum, are interpreted as [m] or [n].

C. Interactive effects of murmurs, transitions, and releases

The perceptual data reported above suggest that:

Either [m] transition structure or [m] murmur structure is a sufficient place cue.

[n] transition structure is a more powerful place cue than [n] murmur structure.

[ɲ] transition structure but not [ɲ] murmur structure is a sufficient place cue.

[ŋ] transitions, murmur, and release are needed for a satisfactory place identification with [æ] but only murmur with [a].

None of the experiments mentioned so far varied all cues simultaneously to assess their relative importance. Moreover, information about their perceptual role was obtained from unnatural acoustic structures, namely isolated murmurs and vowel–consonant sequences with an incomplete set of cues. A more realistic approach is due to Malécot (1956) who compared the cue values of transitions and released murmurs by interchanging them in natural-speech CV and VC utterances with [æ] and [m], [n], [ŋ]; identification was tested with American English subjects. His results are highly compatible with the summary of interactive effects among cues reported above. He found that, for VC syllables, [m] murmur overrode [n], [ŋ] transitions, and [m] transitions overrode [n], [ŋ] murmurs; [n] transitions overrode [ŋ] murmur, and [n] murmur tended to override [ŋ] transitions. According to data summarized above, the effect of [ŋ] murmur would have been presumably much greater in the context of the vowel [a].

The present study used synthetic speech to systematically investigate the role that transitions and murmurs play in the recognition of final unreleased nasals (excluding [m]) following [a] in Catalan, a Romance language spoken in the Eastern part of Spain. Special emphasis was given to the cues for the palatal nasal [ɲ] since their perceptual role was largely unknown. Synthetic transitions and murmurs, with values obtained from real speech utterances, were combined and varied simultaneously in perceptual continua for the three different place categories. The perceptual results will be discussed in relation to production data collected from Catalan speakers and to data on other languages in the literature.

I. ACOUSTIC MEASUREMENTS

A. Method

1. Speakers

Analysis values derived from productions of a single male native speaker of Catalan (the author) served as reference points for the patterns to be used in the synthetic speech experiments. To see to what extent these acoustic patterns were representative, data on the same utterances were subsequently collected from 12 other male native speakers of Catalan.

2. Utterances

Catalan monosyllables ending in VC, V: [a] and C: [m], [n], [ɲ], [ŋ],[2] were read in isolation five times each. Their order was randomized. The final nasal consonants were weakly released or unreleased, depending on the individual speaker. Table I lists the individual utterances in orthographic and phonetic notation; their English meaning is given as well. Subjects were instructed to read the set of utterances as naturally as possible at a normal speech rate. Recordings were made in a soundproof booth; a high-quality microphone and tape recorder were used.

3. Measurements

The utterances were analyzed by means of a Voiceprint sound spectrograph. Readings for murmur spectra ($N1, N2, N3, N4, NZ$) and extents of formant transitions ($F1, F2, F3$) are given in Tables II and III, separately for the reference speaker (means) and for the other 12 speakers (means and standard deviations). Formant frequencies were measured at the center of a formant. Zeros were inferred from frequency areas that showed a major reduction in the magnitude of the energy envelope, as observed in spectral sections. Bandwidth values (bw) for the single speaker were estimated by measuring the distance between two points at the right and left side of the spectrum envelope, equally located 3 dB below the peak level.

TABLE I. Set of utterances used for acoustic analysis.

ham[am] "Aux. we have"	han[an] "Aux. they have"	any[aɲ] "year"	
vam[bam] "Aux. we go"	van[ban] "they go"	bany[baɲ] "bath"	banc[baŋ] "bench"
fam[fam] "hunger"	fan[fan] "they do"		fang[faŋ] "mud"

TABLE II. Analysis frequency values (in Hz) for murmurs in utterances [Cam], [Can], [Caɲ], [Caŋ] produced by a single Catalan subject (means) and 12 other Catalan subjects (means and standard deviations).

	$N1$	$N2$	$N3$	$N4$	NZ	Subjects
[m]	200	1120	1360	2100	···	(1) Single subject
	bw: 160	bw: 120		bw: 180		
	255	1015	1300	1620	···	(2) 12 subjects
	(69)	(69)	(172)	(184)		
[n]	250	850	1550	2025	1780	(1)
	bw: 180		bw: 330	bw: 235		
	285	1035	1515	2130	···	(2)
	(37)	(97)	(62)	(260)		
[ɲ]	250	1025	2100	3125	2650	(1)
	bw: 150	bw: 220	bw: 140	bw: 140		
	295	1055	1760	2265	···	(2)
	(58)	(130)	(355)	(415)		
[ŋ]	350	1200	2030	2540	3700	(1)
	bw: 275	bw: 200	bw: 300	bw: 250		
	295	1060	1530	2215	···	(2)
	(50)	(115)	(145)	(250)		

B. Results and discussion

Murmur values for the reference Catalan speaker (Table II) are consistent with those in the literature. Thus the following trends may be observed:

$N1$ frequency values: [ŋ] > [ɲ],[n] > [m];

$N1$ bandwidth values: [ŋ] > [ɲ],[n],[m];

NZ frequency values: [ŋ] > [ɲ] > [n].

As can be seen in Table II, the data for the single speaker generally fall within two standard deviations from the mean for the other Catalan speakers, except for some large discrepancies in $N3$ and $N4$. This may mean that, in the light of cross-category spectral values obtained from the li-

TABLE III. Analysis frequency values (extent in Hz) for final transitions in utterances [Cam], [Can], [Caɲ], [Caŋ] produced by a single Catalan speaker (means) and 12 other Catalan subjects (means and standard deviations). Negative values indicate falling transitions; positive values indicate rising transitions.

	$F1$	$F2$	$F3$	Subjects
[m]	−120	−285	−10	(1) Single subject
	−70	−75	−55	(2) 12 subjects
	(40)	(93)	(48)	
[n]	−145	+75	+80	(1)
	−30	+235	+65	(2)
	(35)	(97)	(150)	
[ɲ]	−400	+400	+300	(1)
	−120	+460	+125	(2)
	(160)	(125)	(225)	
[ŋ]	−20	+10	−190	(1)
	−30	+110	−65	(2)
	(32)	(138)	(142)	

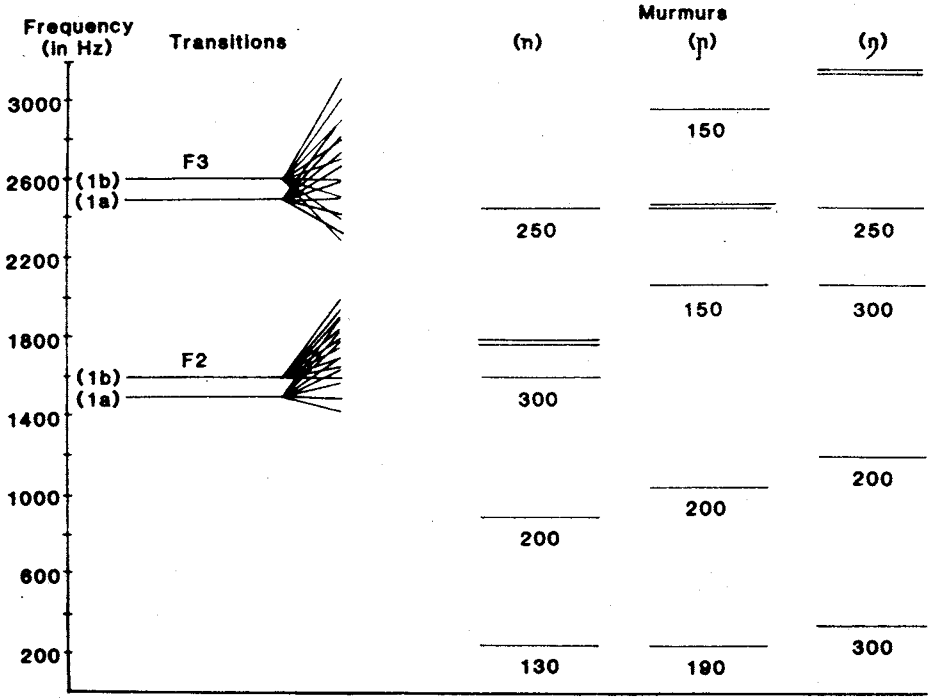

FIG. 1. Synthesis patterns for tests 1a and 1b (fixed murmur conditions) with inclusion of bandwidth values (in Hz) for murmur formants.

terature (see Introduction), $N1$ and NZ but not the upper formants convey relevant information for place identification in Catalan.

The average formant transition frequencies for the single speaker and the group of Catalan speakers (Table III) show the same general directions for $F1$, $F2$, and $F3$ as those reported in the literature (see Introduction). However, $F2$ and $F3$ values for [n] are found to be consistently positive. $F1$ values are strongly negative for [ɲ] and only slightly negative for [ŋ], those for [m] and [n] falling in between. Also, while $F2$ values for [m] (negative) and [ɲ] (positive) fall well apart from those for [n] and [ŋ], those for [ŋ] overlap significantly with those for [n] and [m]. This is consistent with the fact that an appropriate negative $F3$ is needed to synthesize a satisfactory, unambiguous velar nasal consonant (Harris *et al.*, 1958; Fujimura, 1971). The relevance of strongly positive $F2$–$F3$ transition patterns for [ɲ] needs to be emphasized as well (Derkach *et al.*, 1970). In fact, for many speakers such $F2$–$F3$ excursions were still observable during the nasal murmur as an effect of the dynamic motion exhibited by the large mass of tongue body towards the dorsopalatal region. The perceptual relevance of such effects was not investigated in the perception experiment where nasal formants were kept steady during the murmur.

II. PERCEPTION EXPERIMENT

A. Method

1. Stimulus materials

To explore the perceptual role of transitions and unreleased nasal murmurs in place recognition, two types of [an]-[aɲ]-[aŋ] continua were synthesized according to analysis values for the single Catalan speaker (Tables II and III). Considering the complexity of the experimental paradigm, [m] was excluded in order to concentrate on the perceptual role of transitions and murmurs for [ŋ] (rather complex according to data reported in the Introduction) and [ɲ] (unexplored in previous studies). The synthetic stimuli were prepared using a software serial formant synthesizer at Haskins Laboratories (Mattingly *et al.*, 1981) with variable bw parameters, an extra pole (used as $N2$), and an extra zero (used as NZ).

The experiment had two conditions (tests 1 and 2). In test 1 a series of variable $F2$–$F3$ transition endpoints was combined with three fixed murmur patterns believed to be optimal for [n], [ɲ], [ŋ]. In test 2 a series of variable murmur values was combined with three fixed, optimal transition patterns. Schematic spectrograms of the stimuli are presented in Figs. 1 and 2. The poles of the murmur patterns are represented by single lines and the zeros by double lines.

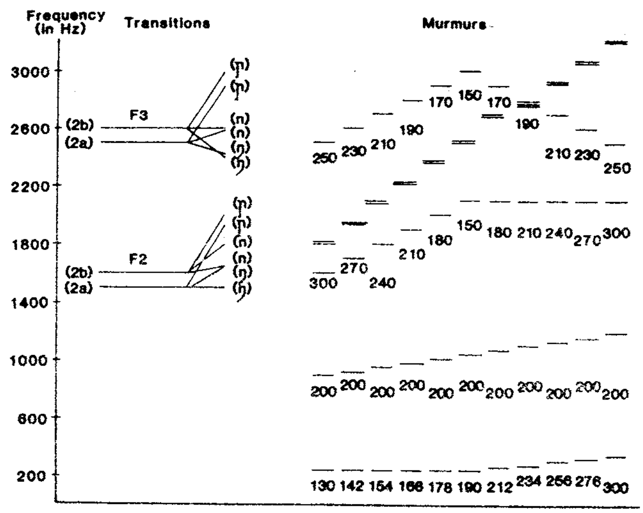

FIG. 2. Synthesis patterns for tests 2a and 2b (fixed transition conditions) with inclusion of bandwidth values (in Hz) for murmur formants.

Two versions of the stimuli (a),(b) were employed which differed slightly in vowel steady state and transition endpoint values. This was done to be able to determine identification cross-over points with higher precision for a larger variety of synthesized $F2$–$F3$ transition values.

Bandwidth values corresponding to the murmur poles were also varied, as indicated in Figs. 1 and 2. $F1$ transitions were flat at 900 Hz in all stimuli. Preceding [a] was nasalized from vowel onset to transition endpoint by means of a progressive frequency rise of a single low pole–zero (from 500/500 to 600/700 Hz) pair. Each stimulus was 560 ms long, having a vowel steady state of 200 ms, a transition of 60 ms, and a murmur portion of 300 ms. There was a progressive 10-dB amplitude decrease from vowel to murmur and a falling $F0$ contour from vowel onset (120 Hz) to murmur offset (80 Hz).

2. Subjects and test procedures

The stimuli were presented in randomized sequences for identification. Test 1a was composed of 144 stimuli (six repetitions of 24 individual stimuli resulting from the combination of eight transition values and three murmur structures), test 1b of 162 (six repetitions, nine transition values, three murmur structures), and tests 2a and 2b of 165 each (five repetitions of 33 individual stimuli resulting from the combination of 11 murmur structures and three transition values). Intervals of 4 s occurred between successive stimuli and longer 10-ms intervals after every ten stimuli. Twenty-four paid Catalan subjects took each of the four tests twice (except for one who did not take test 2). They were asked to identify orthographically the final nasal stop as [n], [ɲ], or [ŋ]. The subjects were all students at the University of Barcelona (Spain) and had had no previous experience with synthetic speech. They did not know the purpose of the experiment. Thirteen of them took the tests binaurally through headphones; the rest listened over a loudspeaker. To check whether the different listening conditions might have had some effect on subjects' responses, the author took all tests under both listening conditions and obtained almost identical cross-over points and response distributions.

TABLE IV. Category judgments (separate percentages for each test and overall percentages for both tests) averaged over continua of transitions (tests 1a, 1b) and murmurs (tests 2a, 2b).

		Response (percent)		
Fixed murmur condition		[ŋ]	[n]	[ɲ]
[ŋ] murmur	test 1a	40.8	26.6	32.6
	test 1b	44.1	23.6	32.2
	Overall	42.4	25.1	32.4
[n] murmur	test 1a	30.3	32.9	36.8
	test 1b	26.7	37.1	36.1
	Overall	28.5	35	36.4
[ɲ] murmur	test 1a	33	31.2	35.8
	test 1b	32.5	31.5	36
	Overall	32.7	31.3	35.9
Fixed transition condition				
[ŋ] transitions	test 2a	63.8	35.3	0.4
	test 2b	53	45.4	1.6
	Overall	58.4	40.3	1
[n] transitions	test 2a	46	51.2	2.8
	test 2b	37.9	54.8	7.1
	Overall	41.9	53	4.9
[ɲ] transitions	test 2a	1.5	2.8	95.7
	test 2b	2.8	4.1	93.1
	Overall	2.1	3.4	94.4

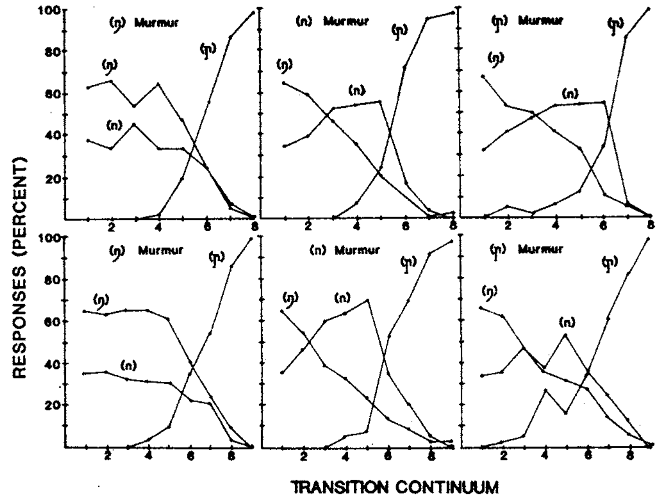

FIG. 3. Perceptual results for tests 1a (above) and 1b (below).

B. Results

Table IV shows category judgments for each test, averaged over the respective continua (transitions in test 1, murmurs in test 2). The detailed results are displayed in Figs. 3 and 4. Each panel in Fig. 3 represents judgments for a particular murmur; the abscissa shows stimulus numbers corresponding to the $F2$–$F3$ transition continuum with increasing endpoint values for [ŋ] < [n] < [ɲ]. Each panel in Fig. 4 represents judgments for a particular $F2$–$F3$ transition set; the abscissa shows stimulus numbers corresponding to the mur-

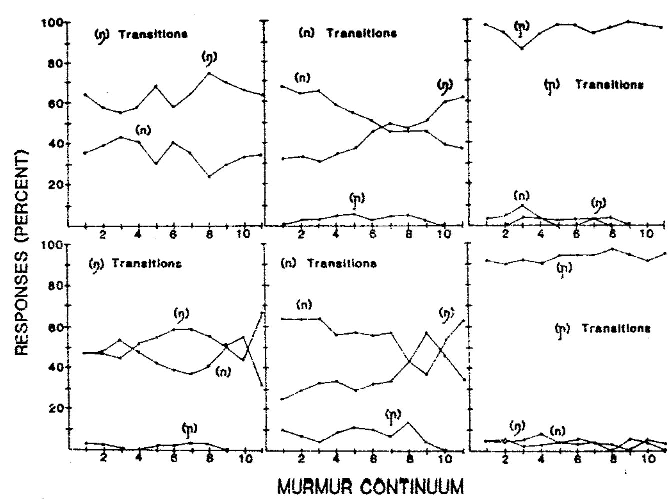

FIG. 4. Perceptual results for tests 2a (above) and 2b (below).

mur continuum with increasing spectral values for [n] < [ɲ] < [ŋ].

Table IV and Fig. 3 provide information about the cue value of optimal murmurs for place identification when attached to appropriate and conflicting transition values. The data show that category judgments cannot be predicted on the basis of appropriate murmur structures and that, except for [n] transitions followed by [ŋ] murmur, optimal transitions largely override conflicting murmurs, completely so in the case of [ɲ] transitions (right end of the transition continuum in Fig. 3). Identification peaks for each place category accord very well with the analysis values for $F2$–$F3$ transitions given in Table III.

Murmurs nevertheless contributed significantly to category identification. Effects of murmurs in test 1 were assessed in analyses of variance of response percentages averaged over all transition values, separately for each response category. The murmur effect was highly significant in the case of [ŋ] responses [$F(2,46) = 15.4, p < 0.001$], and smaller but still significant for [n] [$F(2,46) = 6.2, p < 0.01$] and [ɲ] [$F(2,46) = 4.2, p < 0.05$] responses. [n] murmurs produced somewhat more [n] responses in test 1b than in test 1a, as indicated in a significant murmur-by-test interaction [$F(2,46) = 5.3, p < 0.01$]. Cross-category differences in degrees of statistical significance reveal that the murmur effect was primarily due to [ŋ] murmur versus the other two (see also Fig. 3).

Table IV and Fig. 4 provide information about the cue value of optimal transitions for place identification when attached to a variety of murmurs. The data show again that category judgments tend to be determined by the transitions, particularly in the case of [ɲ] transitions. Except for [n] transitions followed by [ŋ] murmur, optimal murmurs are overridden by conflicting transitions. Again, however, murmurs made a significant contribution. Analysis of variance of [n] responses revealed significant changes across the murmur continuum in the fixed [n] transition condition [$F(10,220) = 6.393, p < 0.001$] and in the fixed [ŋ] transition condition [$F(10,220) = 2.213, p < 0.05$]. Effects between tests 2a and 2b are significantly different for the [ŋ] transition condition [$F(10,220) = 2.863, p < 0.01$] but not for the [n] transition condition. The pattern of murmur effects is predicted by the perceptual relevance of optimal murmur structures; thus, according to Fig. 4, for [ŋ] transitions, [n] responses increase gradually towards an optimal [n] murmur structure (left end of the murmur continuum) in test 2a; for [n] transitions, [ŋ] responses also increase gradually towards an optimal [ŋ] murmur structure (right end of the murmur continuum) in both tests. Murmur effects are stronger in the [n] transition condition than in the [ŋ] transition condition indicating that [n] transitions were a less salient cue than [ŋ] transitions (cf. also Fig. 3).

Neither [n] nor [ŋ] transitions were a totally sufficient cue for place of articulation, even though they had been selected to be optimal for their respective categories; this fact, of course, enabled murmurs to make a contribution in the first place. Optimal [ɲ] transitions, on the other hand, were completely sufficient, so that the murmurs were redundant in this case. However, as Fig. 3 shows, [ɲ] murmurs did not increase the percentage of [ɲ] responses even when the transitions were ambiguous.

Figure 5 compares the perceptual results with the production data gathered from Catalan speakers. The values for the $F2$–$F3$ transitions of the synthetic stimuli of tests 1a and 1b are indicated by crosses lying along the diagonal lines. The ranges of $F2$–$F3$ transitions in productions of [m], [n], [ɲ], [ŋ] are represented by the enclosed areas.[3] The predominant perceptual category judgments which are indicated by the symbols along the diagonal lines are in good agreement with the production spaces for the different nasals, thus confirming the fact that transitions are good place cues. The mismatch between perceptual stimuli and frequency values corresponding to the [ɲ] production space did not affect the quality of [ɲ] judgments, thus suggesting that subjects perceive a positive $F2$ transition as a palatal cue when it points to a critical high locus. It can also be observed that transitions are ambiguous cues for $F2$–$F3$ range values that lie between or on the edges of [ŋ] and [n] production spaces. Thus as shown in Fig. 3, [ŋ] versus [n] murmurs are used by Catalan informants to disambiguate transition patterns with $F2$–$F3$ values that fall at the intersection of both production spaces. This appears to be a good instance of perceptual integration of cues.

C. Discussion

The perceptual results indicate that, in the context of the vowel [a], transitions provide more effective cues for the place of articulation of final nasal consonants than do murmurs. This is consistent with results obtained in previous experiments with synthetic speech. The results also show,

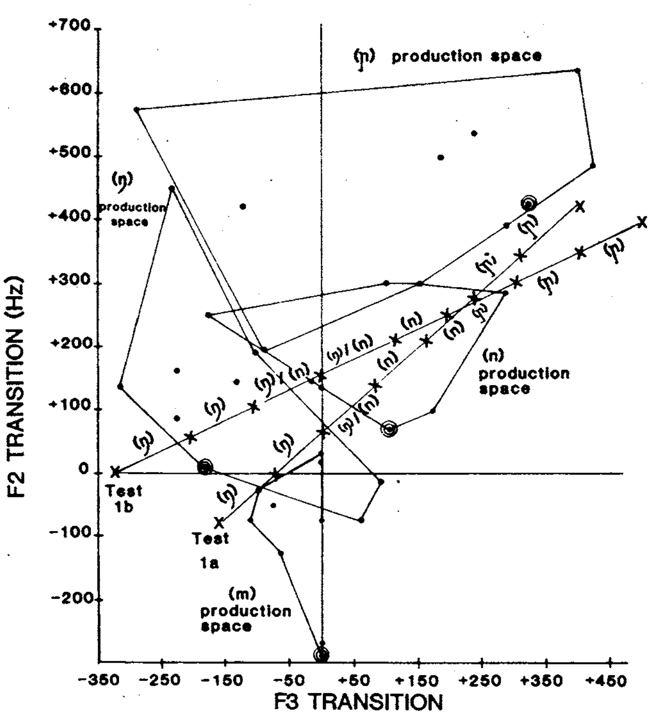

FIG. 5. Comparison of perception and production data with respect to formant transitions. Data for the single Catalan speaker are indicated with dots enclosed by circles.

however, that murmurs contribute significantly to place identification. The design of earlier studies may not have been sufficiently sensitive to reveal these murmur effects.

The present findings about relevance of transitions and murmurs in the identification of different place categories are generally consistent with production and perception data from other languages, summarized in the Introduction. While transitions are generally more powerful cues than murmurs, their relative power differs across place categories. Transitions are highly relevant for [ɲ] but somewhat less so for [ŋ] and [n]. Conversely, [ŋ] and [n] murmurs contribute more to place identification than [ɲ] murmur. Thus there appears to be a reciprocal relationship between the relative salience of the two sets of cues. The finding that [ɲ] murmurs convey very little place information may have been due, in part, to the absence in the synthetic murmurs of the characteristic [ɲ] glide component. Spectrographic analysis reveals that, while very little movement can be detected for nasal formants during the murmur in the case of [n] and [ŋ], that of [ɲ] shows a continuation of positive $F2$–$F3$ transitions, apparently due to a continuation of tongue movement during lingual closure and complete oronasal coupling. On the other hand, the power of the formant transition cues may have been diminished in the present experiment given the fact that contrasting values for $F1$ transitions across different nasal categories were not taken into consideration. This condition, together with the absence of releases in the synthetic stimuli, could perhaps account for the failure to obtain 100% [n] or [ŋ] responses in any of the perceptual tests.

The present results strongly suggest that, in the identification of place for nasals, transitions and murmurs are simultaneously processed in a phonetic mode, i.e., with reference to the relevant products of articulation. However, some of the differences in relative cue salience may also reflect constraints imposed by the auditory system. Psychoacoustic factors may account for the correlation between the extent of the $F2$ transition and its power as a place cue, according to the arrangement [ɲ] > [n], [ŋ]. Klatt and Shattuck (1975) found in experiments with nonspeech stimuli that the perceptual importance of an $F2$-like glide with respect to an $F3$-like glide is positively correlated with its frequency height. Since the $F2$ transition for [ɲ] is much more positive (ends at a much higher frequency) than that for [n] or [ŋ] (cf. Fig. 5), the perceptual data are in agreement with the nonspeech results of Klatt and Shattuck. On the other hand, no auditory constraint is known that would predict the perceptual distinctiveness of [ŋ] versus [n], [ɲ] murmurs.

Thus it appears that certain constraints are imposed by the auditory mechanism before the cues are integrated at a more central stage with reference to articulation. Reference to speech production accounts for the perceptual integration of temporally distributed and spectrally diverse cues (Repp *et al.*, 1978): thus, as shown, transitions and murmurs (and, presumably, releases, had they been present) are evaluated simultaneously and complementarily. Thus the perceptual interpretation of transitions and murmurs for nasals is similar, in principle, to the integration of burst and transitions for non-nasal stops in a CV environment (Dorman *et al.*, 1977) or of fricative noise and formant transitions for prevocalic fricatives (Mann and Repp, 1980; Whalen, 1981). In each case, both relatively static and dynamic cues contribute to a unitary phonetic percept.

ACKNOWLEDGMENTS

This work was supported by NICHD Grant HD01994 and BRS Grant RR05596 to Haskins Laboratories. I acknowledge the assistance of Ignatius Mattingly and Bruno Repp in the preparation of the manuscript. A shorter version of this paper was presented at the 101st Meeting of the Acoustical Society of America in Ottawa, May 1981.

[1]The open vowel [a] has been chosen for analysis since the perception of consonantal nasalization improves with [a] versus [i], [u] (Ali *et al.*, 1971; Mártony, 1964; Zee, 1981). Also [ŋ], which happens to be harder to identify, in general, than other nasal consonants (Malécot, 1956; Ohala, 1975), can be recognized with [a] quite easily when synthesized (Hecker, 1962) or presented in natural speech (Wang and Fillmore, 1961) but rather poorly when it happens to be contiguous to [i], [u].

[2]In Eastern Catalan, the phones [m], [n], [ɲ], [ŋ] appear in absolute final position. [m], [n], [ɲ] also appear intervocalically and correspond to underlying /m/, /n/, /ɲ/. [ŋ] is found word-internally only immediately before [k] and, in final position, occasionally in free variation with [ŋk] depending on the speaker and lexical item. These phonetic facts argue for [ŋ] being an allophone of underlying /n/ before a velar stop.

[3]Such a graph was found to accord more with perceptual processing of nasals than one which relates stimulus points to production data on absolute transition endpoint frequencies: a comparison of results from tests 1a and 1b suggests that, in categorizing stimuli, listeners in fact attended to $F2$–$F3$ transition direction and extent and not to absolute transition endpoints.

Ali, A., Gallagher, T., Goldstein, J., and Daniloff, R. (1971). "Perception of coarticulated nasality," J. Acoust. Soc. Am. 49, 538–540.

Delattre, P. (1958). "Les indices acoustiques de la parole," Phonetica 2, 226–251.

Derkach, M., Fant, G., and de Serpa-Leitão, A. (1970). "Phoneme coarticulation in Russian hard and soft VCV-utterances with voiceless fricatives," Speech Trans. Lab. Q. Prog. Stat. Rep. 2–3 (Royal Institute of Technology, Stockholm), 1–7.

Dorman, M. F., Studdert-Kennedy, M., and Raphael, L. J. (1977). "Stop-consonant recognition: release burst and formant transitions as functionally equivalent, context-dependent cues," Percept. Psychophys. 22, 109–122.

Dukiewicz, L. (1967). *Polskie gloski nosowe. Analiza akustyczna* (Polska Akademia Nauk, Warsaw).

Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Fant, G. (1973). "Acoustic description and classification of phonetic units," in *Speech Sounds and Features* (MIT, Cambridge, MA), pp. 32–83.

Fujimura, O. (1962). "Analysis of nasal consonants," J. Acoust. Soc. Am. 34, 1865–1875.

Fujimura, O. (1971). "Remarks on stop consonants. Synthesis experiments and acoustic cues," in *Form and Substance*, edited by L. L. Hammerich, R. Jakobson, and E. Zwirner (Akademisk Forlag, Denmark), pp. 221–232.

Garcia, E. (1966). "The identification and discrimination of synthetic nasals," Haskins Lab. Stat. Rep. Speech Res. 7–8, 3.1–3.16.

Garcia, E. (1967a). "Labelling of synthetic nasals," Haskins Lab. Stat. Rep. Speech Res. 9, 4.1–4.17.

Garcia, E. (1967b). "Discrimination of three-formant nasal-vowel syllables," Haskins Lab. Stat. Rep. Speech Res. 12, 143–153.

Harris, K. S., Hoffman, H. S., Liberman, A. M., Delattre, P. C., and Cooper, F. S. (1958). "Effect of third-formant transitions on the perception of the voiced stop consonants," J. Acoust. Soc. Am. 50, 122–126.

Hecker, M. H. (1962). "Studies of nasal consonants with an articulatory speech synthesizer," J. Acoust. Soc. Am. 34, 179–188.

Henderson, J. (1978). "On the perception of nasal consonants," Department of Linguistics, University of Connecticut, Storrs, CT (unpublished).

House, A. S. (**1957**). "Analog studies of nasal consonants," J. Speech Hear. Disord. **22**, 190–204.
Klatt, D. H., and Shattuck, S. R. (**1975**). "Perception of brief stimuli that resemble rapid formant transitions," in *Auditory Analysis and Perception of Speech*, edited by G. Fant and M. A. A. Tatham (Academic, New York), pp. 293–301.
Larkey, L., Wald, J., and Strange, W. (**1978**). "Perception of synthetic nasal consonants in initial and final syllable position," Percept. Psychophys. **23**, 299–312.
Liberman, A. M., Delattre, P. C., Cooper, F. S., and Gerstman, L. J. (**1954**). "The role of consonant–vowel transitions in the perception of the stop and nasal consonants," Psychol. Mono. **68**.
Magdics, K. (**1969**). *Studies in the Acoustic Characteristics of Hungarian Speech Sounds* (Indiana University, Uralic and Altaic Series 97).
Malécot, A. (**1956**). "Acoustic cues for nasal consonants: An experimental study involving a tape-splicing technique," Language **32**, 274–284.
Mann, V., and Repp, B. (**1980**). "Influence of vocalic context on perception of the [ʃ] — [s] distinction," Percept. Psychophys. **28**, 213–228.
Mártony, J. (**1964**). "The role of formant amplitudes in synthesis of nasal consonants," Speech Trans. Lab. Q. Prog. Stat. Rep. **3** (Royal Institute of Technology, Stockholm), 28–31.
Mattingly, I., Pollock, S., Levas, A., Scully, W., and Levitt, A. (**1981**). "Software synthesizer for phonetic research," J. Acoust. Soc. Am. Suppl. 1 **69**, S83.
Miller, J. L., and Eimas, P. D. (**1977**). "Studies on the perception of place and manner of articulation; A comparison of the labial-alveolar and nasal-stop distinctions," J. Acoust. Soc. Am. **61**, 835–845.
Nakata, K. (**1959**). "Synthesis and perception of nasal consonants," J. Acoust. Soc. Am. **31**, 661–666.
Ohala, J. J. (**1975**). "Phonetic explanations for nasal sound patterns," in *Nasalfest*, edited by Ch. A. Ferguson, L. M. Hyman, and J. J. Ohala (Department of Linguistics, Stanford University), pp. 289–316.
Repp, B. H., Liberman, A. M., Eccardt, T., and Pesetsky, D. (**1978**). "Perceptual integration of acoustic cues for stop, fricative and affricate manner," J. Exp. Psychol. Hum. Percept. Perform. **4**, 621–637.
Romportl, M. (**1973**). "Zur Akustischen Analyse und Klassifizierung der Nasale," in *Studies in Phonetics* (Mouton, The Hague), pp. 78–83.
Vagges, K., Ferrero, F. E., Caldognetto-Magno, E., and Lavagnoli, C. (**1978**). "Some acoustic characteristics of Italian consonants," J. Ital. Linguist. **3**, 69–85.
Whalen, D. H. (**1981**). "Effects of vocalic formant transitions and vowel quality on the English [s] — [š] boundary," J. Acoust. Soc. Am. **69**, 275–282.
Wang, W. S.-Y., and Fillmore, Ch. J. (**1961**). "Intrinsic cues and consonant perception," J. Speech Hear. Res. **4**, 130–136.
Zee, E. (**1981**). "Effect of vowel quality on perception of post-vocalic nasal consonants in noise," J. Phonet. **9**, 35–48.