

Perception of Alveolar and Velar Allophones of English /l/ in Word-Initial and Word-Final Positions

Author(s): Sara Gesuato

Proceedings of the Twenty-Second Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Role of Learnability in Grammatical Theory (1996), pp. 116-131

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Perception of Alveolar and Velar Allophones of English /l/ in Word-Initial and Word-Final Positions

Sara Gesuato

University of Padua, Italy and University of California at Berkeley

1. Introduction.

From an articulatory point of view, the English /l/ can be defined as a lateral approximant. It is lateral because the airstream is obstructed at some point along the middle of the oral tract, but released along one or both sides of the tongue¹. It is also approximant because the incomplete closure between the side(s) of the tongue and the roof of the mouth narrows the oral tract without producing turbulent airstream (Ladefoged 1993:10).

English /l/ is realized by different allophones depending on which position it occupies in the syllable, namely pre- or post-vocalic².

In articulating a word-initial /l/ the tongue "tip is touching the alveolar ridge, and one or both sides are near the upper side teeth, but not quite touching" and the front of the tongue is raised towards the hard palate (Ladefoged 1993:65). A word-final /l/, instead, is produced by drawing the tongue body up and back in the mouth while optionally maintaining the tongue tip in contact with the alveolar ridge (see Giles and Moll 1975:211-212)³. If the tongue tip is touching the alveolar ridge, the raising of the back of the tongue towards the soft palate is a secondary articulation called velarization. On the other hand, if there is no contact with the alveolar ridge, the sound produced is more like a back vowel or glide (Ladefoged 1993:65; Clark and Yallop 1990:129-130; Hardcastle and Barry 1985:31-33, 41-43)⁴.

These articulatory variations have acoustic correlates. Clark and Yallop (1990:266) explain that if there is a tongue body movement to a posterior position, F_1 will rise modestly and F_2 will drop significantly; also, if there is a tongue body shape in neutral or posterior position, F_2 will fall substantially as constriction increases.

Spectrographic analyses of the positional variants of /l/ can also reveal their acoustic differences. In her study of acoustic variation of American English /l/, Lehiste (1964:14) found that the "main difference between the initial and final allophones of /l/ lies in the relative closeness of F_1 and F_2 in final positions" in the sense that the values for F_1 and F_2 of the final allophone of /l/ are, respectively, higher and lower than those for F_1 and F_2 of the initial allophone of /l/. The average formant positions registered by Lehiste (1964:14) in her control group "were 295, 950, and 2610 cps for initial /l/, and 455, 795 and 2585 cps for final /l/"⁵. Ladefoged (1993:203) also reports very similar formant positions for word-initial /l/, namely 250, 1200 and 2400 Hz⁶.

Both articulatorily and acoustically word-initial [l] resembles [n], while word-final [ɫ] resembles [u] and [w]. Like [l], [n] is pronounced by creating complete obstruction in the mouth in the alveolar region with the tongue tip or blade, while maintaining voicing. However, air does not flow out along the sides of the tongue;

instead, the velum is lowered, which allows the air to go out through the nasal cavity. In the acoustic domain, both [l] and [n] show formant structure similar to that of vowels, but are less clearly marked, and they are characterized by abrupt amplitude and spectral discontinuities at the formant transitions into or out of a vowel (Ladefoged 1993:201, 203, 205; Ohala forthcoming:11). The main difference between the two is that in /l/ there "are but very few and weak harmonics indicating the occurrence of formants ranking at about 2500 cps and characteristics of nasals" (Tarnóczy 1948:5).

[u] is a high back rounded vowel, and, like [ɥ], [u] is a fully voiced sound. It is pronounced by retracting the tongue body towards the back of the mouth and raising it towards the soft palate, and often also by rounding and protruding the lips, which approach each other very closely. Thus the main articulatory difference between [ɥ] and [u] is the presence of lip-rounding in the latter, which further contributes to lowering the value of F_2 (see Clark and Yallop 1990:266).

[ɥ] and [u] are similar also from an acoustic point of view; their formant structures are very much alike⁷, except that [u], being a vowel, has a more intense amplitude than [ɥ].

[w]⁹ is the non-syllabic counterpart — and thus a shorter version — of [u], articulated by creating a labio-velar constriction, which rapidly glides into a vowel. As is the case with any other approximant, its formant structure is similar to that of its corresponding vowel, usually changing (Ladefoged 1993:203), and its articulation varies "slightly depending on the articulation of the following vowel" (Ladefoged 1993:64¹⁰). Thus, [w] is also similar to [ɥ], the difference lying in (a) the longer duration of and (b) the absence of labial constriction in the latter.

The distribution of [l], [n], [w], and [ɥ] in English is subject to a few phonotactic constraints. In syllable-initial position before a vowel, [l], [n] or [w] can occur, but the velarized [ɥ] cannot, although [ɥ] is articulatorily and acoustically similar to [w]. In word-final position after a vowel, [w]¹¹, [ɥ] or [n] can occur, but [l] cannot, although it is articulatorily and acoustically similar to [n]. The sounds [w] and [n] are thus similar to, but only in partially overlapping distribution with, the sounds [ɥ] and [l], respectively. As a result, they belong to different phonemes, [l] and [ɥ] being subsumed under /l/, [n] under /n/ and [w] under /w/.

Given that [l] and [ɥ] are only partially similar to each other and given that they cannot replace each other in all environments, it is likely that they are recognized as allophones of the same phoneme /l/ not only because of their physical properties, but especially because of their distributional constraints.

It is thus possible to hypothesize that if an alveolar [l] should occur word-finally, and if a velar [ɥ] should occur word-initially, they might be perceived as /n/ and /w/, respectively, given that these sounds are similar to the /l/ allophones, but without being limited in their distribution. In the case of a syllable-final alveolar [l], misperception should occur more frequently in the presence of a preceding front vowel, given that a more velarized variant of /l/ is usually realized in this environment. Hardcastle and Barry (1985:43) note that in some dialects of British English, velarized [ɥ] tends to be realized as a non-syllabic back vocoid¹² "more frequently with front vowels than with back vowels," and they explain this "within a perceptual framework. The velar component of" [ɥ] "contrasts more clearly with

front vowels than back vowels making the contribution of actual alveolar contact for the /l/ identification less important." If this is the case, then, a native listener of English is likely to be especially thrown off by a light [l] following a front vowel¹³.

An experiment was designed to determine whether and to what extent the light and dark allophones of /l/ can be misperceived as /n/ and /w/, respectively, depending on where they occur in the word and on the surrounding vocalic environment.

2. Experimental design.

2.1. Subjects.

Twenty native and three non-native¹⁴ speakers of American English (fifteen males and eight females) served as volunteer subjects for this experiment. They were not aware of the purpose of the experiment, nor had they received any phonetic training¹⁵.

2.2. Stimuli.

It was decided to investigate perception of /l/ allophones in two environments: (a) in pre-vocalic position in /lVC sequences, and (b) in post-vocalic position in CV/l/ sequences. Language material was obtained from the speech of two female graduate students in linguistics at UC Berkeley, an Italian speaker (this author) and an American speaker.

Given that the quality of /l/ articulations depends on their surrounding phonetic context (see footnote 3), the speech sample was selected so as to exemplify /l/ in a variety of pre- and post-vocalic contexts. In the Italian words, V was varied over each of the attested Italian vowel phonemes /i e ε a o u/. In the English words, it was varied over most of the English vowel phonemes, namely /i: ɪ æ ə-¹⁶ ʌ ɒ: ʊ u/¹⁷.

As for the /l/, the Italian speaker was the only one who recorded words with word-final alveolar [l] — the only type of /l/ to be found in Italian — and the English speaker was the only one who recorded words with word-final [ɫ], given that their competence on their respective native languages enabled them to accurately pronounce those sounds in those environments.

The non-/l/ consonant, which was not under investigation, took a value over a small subset of Italian and English consonants (e.g., Italian [f i l] *fil* 'thread'; English [s i l] *sill*; see Table 1 for a complete list). In the tokens to be reversed (see below), this C was always a continuant so as to ensure that, even when played backwards, these words would still sound natural. Whenever possible, CVC sequences were selected that corresponded to real English words. To this end, at first only English words were recorded. However, since it was hard for the Italian speaker to pronounce word-final alveolar [l] in *English* words, it was then decided that she should record Italian words instead.

Recordings were made on a CSL system using a high quality microphone at a cut-off frequency of Hz 20,000. The CVC words that were not going to be played backwards were recorded in the frame sentence 'Say X for me', X being the target

word, so that all the target words could equally receive primary stress and highest relative intonation level. The CVC sequences to be reversed¹⁸, instead, were recorded in the frame 'Will you say X?': that is, they were pronounced with a rising intonation to ensure that, once reversed, they would sound natural, and similar to the sets of stimuli that were not reversed¹⁹.

A white noise waveform was added to the stimuli²⁰. In each case, it was the same duration as, but 24 dB lower than, the peak loudness of the loudest signal in the CVC word it was superimposed on. The presence of white noise was supposed to increase the likelihood of the subjects' misperception of word-final [l] as /n/ and of word-initial [ʃ] as /w/.

The speech sample consisted of both English and Italian CVC syllables (see Table 1 below). They were partly real and partly nonsense words, which, however, respected the phonotactics of the two languages. Sample set I was used to study perception of word-final [l], and sample set II was used to study perception of word-initial [ʃ].

Sample set I consisted of seven Italian words ending in [l] (column A: [f i l] *fil* 'thread'; [d e l] *del* 'of the (Mac. Sing.)'; [ʃ i e l] *ciel* 'sky, heaven'; [m a l] *mal* 'bad, evil'; [g o l] *gol* 'goal (in soccer)'; [s o l] *sol* 'sun'; [m u l] *mul* 'donkey'), seven Italian words ending in [n] (column B: [v i n] *vin* 'wine'; [m e n] *men* 'less, fewer'; [b e n] *ben* 'well, good'; [m a n] *man* 'hand'; [f o n] *fon* 'hairdryer'; [k o n] *con* 'with'; [s u n], a made-up word)²¹, four English words ending in [ʃ] (column C: [b e ʃ] *bell*; [k æ ʃ] *Cal*; [p o ʃ] *Paul*; [t u ʃ] *tool*), nine made-up words respecting English phonotactics, which were originally spoken as beginning with [l] (i.e., [l i m]; [l i s]; [l e f]; [l æ v]; [l ʌ n]; [l o s]; [l u θ]; [l u z]; [l ə ʃ]), but which were presented to listeners reversed (column D: [m i l] *meal*; [s i l] *sill*; [f e l] *fell*; [v æ l], a made-up word; [n ʌ l] *null*; [s o l] *Saul*; [θ u l], a made-up word; [z u l], a made-up word; [t ə l], a made-up word), the first eight words from column D but played in the original sequence (column E: [l i m], a made-up word; [l i s], a made-up word; [l e f], a made-up word; [l æ v], a made-up word; [l ʌ n], a made-up word; [l o s] *loss*; [l u θ], a made-up word; [l u z] *lose*), and 4 English words ending in [n] reversed from original [n]VC sequences (column F: [s i n] *seen*; [θ i n] *thin*; [ʒ o n], a made-up word; [s u n] *soon*). Except for the third subset (column C), all the above tokens were recorded by the Italian speaker.

Apart from the words in the first subset (see column A), all the others served as control stimuli, that is, they were included in order to assess the well-foundedness of the experiment's hypothesis, and the reliability of the subjects' perception of familiar types of CVC words vis-à-vis that of the unfamiliar CV[l] and [ʃ]VC words (see below).

2.3. Hypothesis.

It was expected that the Italian CV[l] words listed in column A would be perceived as CV/n/, especially after a front vowel, for the reasons outlined in Section 1 above.

CV[n] and CV[ʃ] stimuli (columns B, C, and F) were expected to be correctly identified, since they exemplified types of CVC sequences that are familiar to

English speakers, even if some of them had been reversed and uttered with an Italian accent. They were thus supposed to test the subjects' reliability as listeners.

CV[l] stimuli (columns A and D) — some of which also had been reversed and pronounced with an Italian accent — were instead expected to be misidentified because they consisted of unfamiliar sound sequences.

Items in column D, like those in column A, were not only included to test the hypothesis that misperception of word-final /l/ uttered by an Italian speaker should be greater than misperception of word-final velar [ʃ] uttered by an American speaker (see items in column C), but also to make sure that American listeners' misperception of alveolar [l] as /n/ in a word uttered by an Italian speaker would be greater if [l] occurred word-finally rather than word-initially (see items in column E). Thus the items in columns D and E were supposed to show that an alveolar [l] (the only type of /l/ to be found in Italian) would not be misperceived word-initially, even if pronounced by an Italian.

These words were also used to check whether the rate of misperception could be affected (i.e., increased) by the fact that they had been reversed, since they might sound a bit distorted around the transitions²² and due to the unusual length of a word-final /l/²³.

Sample set II consisted of reversed CV[ʃ] and CV[n] sequences uttered by the American speaker ([ʃ i n]; [ʃ e s]; [ʃ æ ʒ]; [ʃ ʌ ʃ]; [ʃ ɒ z]; [ʃ u θ]; [ʃ u n]; [ʃ ə l]; [n i f]; [n æ ʃ] *gnash*; [n ə ð]; [n u m]; [n ə v] *nerve*). It was expected that the reversed CV[ʃ] sequences would be misperceived as /w/VC (see Section 1 above), but that the reversed CV[n] sequences would be correctly identified as /n/VC words. The latter were included in the stimuli set in order to check, once again, the reliability of the subjects' perception. Original utterances of [ʃ]VC sequences were not included in this set of tokens, simply because it was not possible to find a speaker who could reliably pronounce [ʃ] word-initially²⁴.

The final set of CVC stimuli thus obtained consisted of seven Italian and nine English words ending in [l] (the latter reversed from the original [l]VC sequence in which they had been uttered), four English words ending in [ʃ], eleven words ending in [n] (seven of which were Italian words and four of which were reversed English words), nine English words beginning with [ʃ] (reversed from the original utterance), and five English words beginning with [n] (also reversed from the original).

Table 1 reproduces all the stimuli used in the experiment, listed according to the height of the syllable nucleus²⁵.

The stimuli were divided into three groups, namely (a) words ending in -l or -N (see Table 1, sample set I, columns A, B, C, D and F); (b) words beginning with L- or W- (Table 1, sample set II, column G); and (c) words beginning with L- or N- (Table 1, sample set I, column E, and sample set II, column H), and each sequence was randomized independently. This was done to make the experiment easier for the subjects, so that within each set of stimuli their attention would be focused on only one type of task at a time.

Subjects were informed that they would hear three randomized sets of both real and nonsense words uttered by two female speakers. For each stimulus that was played to them, they would have to identify which one of the two choices on their answer sheet corresponded to the initial (second set) or final (first and third

Table 1. List of stimuli grouped according to phonetic environment, (i.e., height of syllable nucleus), followed by expected (left-hand side) and unexpected (right-hand side) target responses.

Note: *Am.* and *It.* stand for 'uttered by the American speaker' and 'uttered by the Italian speaker,' respectively; *rev* stands for 'reversed from the original utterance.'

Set I. Stimuli for testing perception of word-final [l]						
syllable nucleus	A. It	B. It	C. Am.	D. It (rev)	E. It	F. It (rev)
[i]	1. [f i l] NL	8. [v i n] NL		19. [m i : l] NL	28. [l i : m] LN	36. [s i : n] NL
[ɪ]				20. [s ɪ l] NL	29. [l ɪ s] LN	37. [θ ɪ n] NL
[e]	2. [d e l] NL	9. [m e n] NL				
[ɛ]	3. [tʃ ɛ l] NL	10. [b ɛ n] NL	15. [b ɛ tʃ] LN	21. [f ɛ l] NL	30. [l ɛ f] LN	
[æ]			16. [k æ tʃ] LN	22. [v æ l] NL	31. [l æ v] LN	
[ʌ]				23. [n ʌ l] NL	32. [l ʌ n] LN	
[ɑ]	4. [m ɑ l] NL	11. [m ɑ n] NL				
[ɒ]			17. [p ɒ tʃ] LN	24. [s ɒ : l] NL	33. [l ɒ : s] LN	38. [ʒ ɒ : n] NL
[ɔ]	5. [g ɔ l] NL	12. [f ɔ n] NL				
[o]	6. [s o l] NL	13. [k o n] NL				
[ʊ]				25. [θ ʊ l] NL	34. [l ʊ θ] LN	
[u]	7. [m u l] NL	14. [s u n] NL	18. [t u tʃ] LN	26. [z u : l] NL	35. [l u : z] LN	39. [s u : n] NL
[ə]				27. [t ə l] NL		

Set II: Stimuli for testing perception of word-initial [t]

syllable nucleus	G - Am (rev)	H - Am (rev)
[i]	40. [t i: v] W L	
[ɪ]	41. [t ɪ n] W L	49. [n ɪ f] N L
[e]	42. [t e s] W L	
[æ]	43. [t æ ʒ] W L	50. [n æ ʃ] N L
[ʌ]	44. [t ʌ ʃ] W L	
[ɑ]	45. [t ɒ z] W L	51. [n ɑ ð] N L
[ʊ]	46. [t ʊ θ] W L	
[u]	47. [t u: n] W L	52. [n u: m] N L
[ə]	48. [t ə l] W L	53. [n ə v] N L

set) sound of the word they had just heard. For words in the first and third sets, the choice was between L and N; for those in the second set, it was between L and W.

In the pilot study, run with three people, subjects were also instructed to check the word 'insecure', typed on a separate column on the answer sheet, if they felt very insecure about the answer they had chosen. However, since none of the subjects took advantage of this option, it was then decided to exclude it from the rest of the study.

The stimuli were played through the earphones in the same randomized sequence to all subjects in individual sessions.

Table 2 shows the list of words used in the training session, followed by the answers subjects had to choose from. Subjects had to pay attention to the final segments in the first four words, and to the final segments in the last four words.

The experimental procedure was explained in detail, repeatedly if necessary, and all questions answered, except no clue was given as to what the right or expected answer should be. Each stimulus was played three times²⁶ with a 500-msec interval before each repetition, and separated from the next stimulus by a three-second interval. A brief training session was administered to the subjects, which followed the same procedures as outlined above for the experiment. The reason for this was to make sure that subjects understood what was required of them.

Table 2. Words played during the training session, followed by expected (left-hand side) and unexpected (right-hand side) target responses.

Stimuli exemplifying word-final [l], [n], and [t]			
1. [n e l] N L	2. [n o n] N L	3. [p æ t] L N	4. [f i l] N L
Stimuli exemplifying word-initial [l], [n], reversed [l], and reversed [n]			
5. [l i f] L N	6. [n u z] N L	7. [n ʌ f] N L	8. [t ʌ n] W L

3. Results.

All subjects were considered reliable listeners, since they correctly identified all or most (see below) of the tokens which consisted of familiar types of sound sequences (see Table 1, except for columns A and G). Only two subjects mistook word-initial [n] for /l/ in item # 52, and only one mistook word-final [l] for /n/ in item # 29.

All subjects but two said that they considered the experimental task to be easy for the *most* part (i.e., except for the words listed in Table 1, Set II; see below), and thus added that for the most part they avoided random guessing. They also thought that white noise mixed with the signal was the unavoidable hiss of the recording equipment, and were able to tune it out.

As expected, familiar CVC sequences in the training session were correctly identified, while two of the three unfamiliar ones (i.e., Table 2, #s 1 and 8, but not # 4) were misperceived, probably also because of their unexpectedness. The unfamiliar word # 4, instead — which was # 5 played backwards — was not misunderstood; this was probably due to the duration of its /l/, longer than the one in item # 1 (see note # 22 above)²⁷.

An overall comparison between the control stimuli (i.e., columns B, C, E, F, and H vs. the experimental stimuli (i.e., columns A, D, and G) reveals that the data obtained from the two groups of stimuli are significantly different (chi-square test: $F(1) = 340.48252$, $p < 0.05$).

Inter-set comparison shows that misperception was more marked where it was expected and almost non-existent where it was not. That is, on the one hand, the

rate of misperception of word-final [l] as /n/ (a) in Italian words pronounced by the Italian speaker (column A) and (b) in reversed English words pronounced by the Italian speaker (column D) was significantly different from the rate of misperception of:

(i) word-final [n] as /l/ in Italian words pronounced by the Italian speaker (column B; chi-square test: $F(1) = 18$, $p < 0.05$ and $F(1) = 57$, $p < 0.05$, respectively);

(ii) word-initial [l] as /n/ in English words pronounced by the Italian speaker (column E; chi-square test: $F(1) = 15.21052$, $p < 0.05$ and $F(1) = 54.06896$, $p < 0.05$, respectively);

(iii) word-final [t] as /n/ in English words pronounced by the American speaker (column C; chi-square test: $F(1) = 18$, $p < 0.05$ and $F(1) = 57$, $p < 0.05$, respectively);

(iv) word-final [n] as /l/ in reversed English words pronounced by the Italian speaker (column F; chi-square test: $F(1) = 18$, $p < 0.05$ and $F(1) = 57$, $p < 0.05$, respectively).

On the other hand, the rate of misperception of word-initial [t] as /w/ in reversed English words pronounced by the American speaker (column G) was significantly different from the rate of misperception of:

(a) word-final [t] as /n/ in English words pronounced by the American speaker (column C; chi-square test: $F(1) = 175$, $p < 0.05$); and

(b) word-final [n] as /l/ in reversed English words pronounced by the Italian speaker (column F; chi-square test: $F(1) = 175$, $p < 0.05$)

The above analysis, therefore, shows that the data gathered cannot be said to be due to chance. However, the data do not perfectly match the original expectations either.

Fig. 1 shows the number of misidentifications of word-final [l] as /n/ in CV[l] stimuli (vertical axis) as a function of the syllable nucleus (horizontal axis) in Italian syllables (white columns) and inverted English syllables (shaded columns).

Unlike what was expected, misperception of word-final [l] as /n/ occurred only occasionally, namely 19.90% of the time. However, as expected, it didn't occur at random, and this rate of error was significant in comparison to the listeners' misperception of word-initial [l] (see paragraph (ii) above). In the Italian words, it occurred especially when the syllable nucleus was a high front vowel (chi-square test: $F(1) = 14.22222$, $p < 0.05$).

As Hardcastle and Barry (1985:43) point out, this is the environment which is most likely to trigger — and thus where listeners expect to hear — a vocalized version of a back velar lateral [ɫ]. A velar [ɫ] contrasts more clearly with front vowels, without jeopardizing the hearers' identification of it as /l/ (see Section 2.3 above). Also, after a front vowel, a dark [ɫ] is less likely to be confused with the final glide of a diphthong, because English doesn't have diphthongs consisting of a high front vowel and a labio-velar glide. Thus, subjects misunderstood [l] for /n/ in those environments where they unconsciously were expecting either a very dark [ɫ] or not a lateral at all. Notice, however, that the syllable nucleus that favored the highest number of misperception was [e], that is, not the highest vowel of all.

The rate of misperception was greater among the reversed English words. The reason for this may simply be the distortion caused by the stimuli being played

in reverse, as well as the unexpected longer duration of these [l]s in word-final position. However, contrary to my expectations, the number of misperceptions of [l] as /n/ after front vowels (i.e., [i:], [i], [e], and [æ]) was not significantly higher than the number of misperceptions after back vowels (i.e., [o], [u], and [u:]; chi-square test: $F(1) = 0.1915$, $p < 0.01$).

Also, misperception occurred more frequently in the case of stimuli having non-low vowels rather than low vowels as their syllable nuclei (chi-square test: $F(1) = 19.10526$, $p < 0.01$), which was not expected.

It may be that anticipatory articulation is responsible for this (see Gartenberg 1984: 145-154): that is, it is possible that I uttered these [l]s with a wider (or temporally longer) tongue-palate contact in anticipation of the following high vowel.

At the end of the experiment, the subjects said that they had had no difficulties in deciding between L and N when they had to identify the initial or final segments of these stimuli. When I debriefed them, they were surprised to learn that they had misunderstood some word-final [l]s as /n/s, their most frequent comment being "It really sounded like an N." Thus, although misidentification was not as high as expected, and elicited through a forced choice between only two alternatives, in the cases when it did occur, it revealed what these subjects really had perceived.

However, two subjects stated that the final segment in word # 20 (i.e., [s i l], reversed from [l i s]) sounded like M, and another said the same about word # 27 (i.e., [t ə l], reversed from [l ə t]). They also explained that, since neither N nor L was an adequate alternative for them, they had had to choose their answer at random (# 20 was classified once as L and once as N, and # 27 as L). As far as I know these were the only cases within this group of words in which guessing occurred; however, I do not know how many other people experienced the same perception with those stimuli, but failed to express it to me.

Anticipatory and perseverative coarticulation may be responsible for this unusual type of misperception. In the case of [s i l] — which was later reversed to [l i s] — it is possible that (a) in anticipation of the following front vowel and (b) due to the Italian speaker's accent, the [l] was articulated in the dental, not the alveolar region; since there was no immediately surrounding context that could have triggered (and thus justified) a dental articulation (as would be the case, instead, of the *-n-* in *tenth*, for example), the place of articulation could be reinterpreted as [+labial].

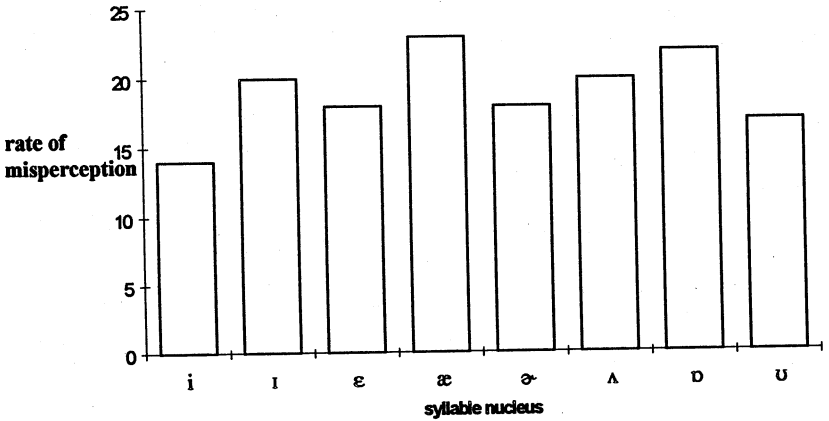
In the case of [t ə l] — reversed from [l ə t] — it is possible that the rhotacization of the syllable nucleus contributed to lowering the F_2 of the preceding segment, and since low F_2 can be induced by labial constriction as well, the acoustic properties of [l] were probably ascribed to an inferred labial articulation²⁸.

Fig. 2 shows the percentage of identification of word-initial [t] as /w/ in [t]VC stimuli (on the vertical axis), as a function of the vowel quality of the syllable nuclei (on the horizontal axis). As expected, most word-initial [t] were mistaken for /w/²⁹. Inter-subset comparison shows that the rate of misperception of word-initial [t] as /w/ was significantly different from the rate of misperception of word-final [t] as /n/ ($F(1) = 175$, $p < 0.05$) as well as from the rate of misperception of

Fig 1. Rate of subjects' misidentification of word-final alveolar [l] as /n/ in CV[l] stimuli (vertical axis) as a function of the syllable nucleus (horizontal axis) in Italian syllables (white columns) and inverted English syllables (shaded columns).



Fig 2. Rate of subjects' misidentification of word-initial [ʃ] as /w/ in [ʃ]VC stimuli (vertical axis) as a function of the vowel quality of the syllable (horizontal axis).



word-initial [n] as /l/ in reversed English words (chi-square test: $F(1) = 169.09038$, $p < 0.05$).

The quality of the syllable nucleus didn't affect the responses in any noticeable way. Within this set, the difference between misperception in words having front vowels and words having back vowels as their nuclei was insignificant (chi-square test: $F(1) = 1.91044$, $p > 0.05$), as was the difference between words having high vowels and words having back vowels as their nuclei (chi-square test: $F(1) = 0.64102$, $p > 0.01$).

Despite their agreement in their responses (on average, /w/ was the preferred answer over /l/ 84.53% of the time), several subjects complained about the difficulty in choosing between L and W in one or more of the first four [ʔ]VC stimuli they heard (namely, #s 41, 43, 46, and 47 in Table 1), and a few said that they had to make a guess in those cases. Nobody complained about the other [ʔ]VC words, but maybe this is because the subjects had gotten used to that type of stimuli by then. The difficulty probably arose out of two flaws in the stimuli: (a) distortion due to the fact that they were played in reverse, and (b) lack of labial coarticulation in the velar [ʔ].

Given that for the most part the subjects did not complain in a similar way about the choice between L and N in the other sets (see above), it means that they were more aware of the distinction between [l] and [n] than between [l] and [ʔ] in their respective environments. This is nicely exemplified in the word [ʔə l], whose word-final [l] was misunderstood as /n/ 49.13% of the time (item # 27), but whose word-initial [ʔ] was misunderstood as /w/ 95.65% of the time (item # 48).

The data gathered thus support the hypothesis that word-final and word-initial positions are important cues for the accurate perception of [ʔ] and [l], respectively, as allophones of /l/.

4. Conclusion.

At issue in this experiment was the assessment of English speakers' ability to discriminate between velar [ʔ] and [w] word-finally, and between [l] and [n] word-initially. Because of the distributional constraints affecting [ʔ] and [l], and their phonetic similarity to [w] and [n], respectively, it was expected that [ʔ]VC syllables would be perceived as [w]VC words, and that CV[l] syllables would be perceived as CV[n] words.

In the experiment outlined above, subjects who could correctly identify word-initial [l] and word-final [ʔ] as allophones of /l/ encountered some difficulties in recognizing those sounds as allophones of the same phoneme if they appeared in unusual positions. As expected, word-initial [ʔ] was usually perceived as /w/, independently of the quality of the following vowel. Instead, unlike what was expected, word-final [l] was often accurately perceived as /l/. However, when misperception did occur, it appeared to be influenced (a) by the frontness of the preceding vowel — as expected — among the Italian stimuli, and (b) by the height of the preceding vowel — contrary to the original hypothesis — among the reversed English words. It was thus suggested that in case (a) subjects failed to recognize [l] as /l/ in those contexts in which, if an /l/ appears, it is (expected to be) realized as a very dark variant (i.e., after a front vowel); in case (b) subjects

were more sensitive to a hypothesized longer duration or wider extension of tongue-palato contact in the alveolar region due to anticipatory coarticulation.

At least two of the English CV[l] stimuli (i.e., [s ɪ l] and [t ə l]) were perceived as CV/m/ words by a few subjects, that is, the potentially ambiguous word-final segment [l] was misperceived both in its manner and in its place of articulation (i.e., nasal instead of liquid and bilabial instead of alveolar, respectively). This phenomenon was tentatively accounted for by means of anticipatory and perseverative coarticulation (see Section 3 above).

The rate of misperception of [l] as /n/ and of [t] as /w/ shows the extent to which ascription of sounds to certain phonemic categories depends on one's (a) intuitive knowledge of what are the expected sound sequences in one's language and (b) sensitivity to the effects of the surrounding environment, rather than on (c) the intrinsic phonetic (articulatory and acoustic) properties of the actual allophones being heard.

This study, therefore, gives support to the claim that grouping of allophones under the same phonemic heading is determined by "a combination of gross phonetic detail plus distributional information" (Ohala 1986:22).

NOTES

¹ The amount of lateral tongue-palate contact depends "on the vowel environment. When following an open back vowel there is usually little or no side contact, but when following a close front vowel there may be almost complete bilateral contact." (Hardcastle and Barry 1985:33).

² Additionally, allophonic variation in /l/ also depends on (a) rate of utterance, (b) position of /l/ in the morpheme, word or phrase, and (c) whether the /l/ is consonantal or syllabic (Lehiste 1964:10; Giles and Moll 1975; Bladon and Al-Bamerni 1976; Gartenberg 1984; Sproat and Fujimura 1994:292-293). At the same time, the different allophones of /l/ affect the quality of surrounding vowels. For example, "syllable-final /l/ in *mealy* has an influence on the preceding vowel, lowering it — this does not occur in freely, where /l/ is not morpheme-final." (Arthur Bronstein, December 13, 1995: personal communication). In this study, I am concerned only with perception of pre- and post-vocalic allophones of non-syllabic /l/. Given the type of stimuli used in the experiment described below, the terms pre-vocalic, syllable-initial, and word-initial on the one hand, and the terms post-vocalic, syllable-final, and word-final on the other will be used interchangeably throughout the paper.

³ Sproat and Fujimura (1993) derive the phonetic differences between light and dark /l/ partly from durational considerations ((a) the light /l/ and the syllable in which it occurs have a shorter duration than dark /l/ and thus the tongue retraction does not have the time to reach its full target; and (b) "the value of the lag between the time at which the tongue tip reaches its extremum of forward movement and the time at which the tongue tip reaches its extremum of retraction and lowering movement [...] is greater for dark /l/s than for light /l/s [...]") (pp. 298-300), and partly from expected differences between pre- and post-vocalic consonants: they propose that all /l/s involve an apical and a dorsal gesture. The former is consonantal, while the latter is vocalic, since it does not produce a radical constriction in the vocal tract. They also assume that "consonantal gestures tend to be stronger (i.e., have greater displacements) in syllable-initial position and weaker in syllable-final position," and that the opposite holds for vocalic gestures. They relate these ideas "to the fact that the universally basic syllable type is CV: consonantal gestures are more typically manifested at the beginning of syllables and vocalic gestures at the end," and point out that "positing two distinct allophones for English /l/ has the effect of obscuring the broader generalizations that are true not just of syllable-initial and syllable-final /l/, but syllable-initial and syllable-final consonants of other types too; [...]" (pp. 304-305).

⁴ Spectrographic analyses (Bladon and Al-Bamerni 1976; Lehiste 1964) and electro-palatographic studies (Gartenberg 1984; Hardcastle and Barry 1985) of English /l/ articulations also revealed that both allophones are in turn subject to variation in quality depending on the surrounding phonetic context. In a cinefluorographic study of allophonic variations of English /l/, Giles and Moll (1975:210-211) noticed that "for each of the subjects the tongue dorsum assumes approximately a constant shape for all of the /l/ productions; however, the lingual apex and root show variations in contour." By comparing dorsum contours for /l/ "with contours selected from steady state portions of the different vowel articulations" they also showed that "the lingual dorsum shape assumed during production of /l/ allophones is essentially the same as that observed during production of the various vowels studied." (For comments on the influence of vowel context on perception of word-final alveolar [l] see Section 3 below.)

⁵ There is also considerable interplay between allophonic variation of /l/ and variation of the surrounding phonetic context. In analyzing vowel allophones associated with initial and final /l/, Lehiste (1964:10) points out "[...] that the second formant of the initial allophone of /l/ anticipates to a certain degree the second formant position of the following vowel, but that the first and third formants of the vowel are in turn influenced by the preceding /l/." That is, the first formant of the syllable nucleus following an initial /l/ "is, as a rule, higher than the average value for the first formant" (p. 23). She adds that this supports Fant's findings that "the recognition of /l/ is accomplished in a considerable degree by the rapid shift up of F₁ from the lateral to the following vowel" (p. 23). Lehiste also states that the third formant for an initial /l/ is also noticeably higher than the average (p. 23). On the other hand, "the final allophone of /l/ shows a much smaller range of variations and is essentially independent of the preceding vowel, but it exerts a strong influence on the second formant of the preceding syllable nucleus" (p. 10), that is, "when the vowel is followed by /l/, the second formant of the vowel is always considerably lower than the average" (p. 26).

⁶ Bladon and Al-Bamerni (1976:141-143) report their frequency measurements in mels on three charts from which it is impossible to compute the exact mean values for pre- and post-vocalic /l/. However, in their discussion (p. 142) they state that the mean value of F₂ for word-initial /l/ is 1218 mels.

⁷ Formant values for [u] given by Ladefoged (1993:193) are F₁ 310, F₂ 870, and F₃ 2250; those for [ɪ] reported by Lehiste were given above in this section.

⁸ Lehiste (1964:32) notes that it is "the formant position of the syllabic [emphasis mine] allophone of /l/" that "resembles that of the back vowel /u/ [...]"

⁹ In this paper I use the symbol [w] to stand both for the syllable-initial labio-velar onglide and for the syllable final labio-velar offglide.

¹⁰ Lehiste (1964:122) reports average formant values for [w] as second element of /ou/ and /au/ diphthongs which are very similar to those quoted above for velar [ɪ] and for [u], namely F₁ 410-545, F₂ 740-870, and F₃ 2225-2335.

¹¹ In this paper [w] is taken to represent also the off-glide at the end of such diphthongs as are to be found, for example, in the word *how*.

¹² Arthur Bronstein (personal communication, December 13, 1995), however, observes that "typically, British English /l/ tends to be 'clearer,' in final position, than American English /l/." For a detailed description of the range of /l/ realizations in different varieties, see Wells 1982: 258-9, 313-317, et passim).

¹³ The author's personal experience also tends to give support to this hypothesis. Being an Italian speaker, she often mispronounces the letter *L* as [ɛ l] when spelling names, and English speakers mistake her [ɛ l]s for [ɛ n]s.

¹⁴ The latter were first-generation immigrants to the US, who had been exposed to American English only since kindergarten. None of these non-native speakers were still fluent in their native languages. All were, however, fluent in English.

¹⁵ It was later discovered that one subject was a polyglot and a certified translator. She had received phonetic training in languages other than English, especially Polish.

¹⁶ In this paper no distinction is made between stressed [ɜ̄] and unstressed [ə̄].

17 Since some of the tokens had to be played backwards (see below), the syllable nucleus never consisted of a diphthong, so as to avoid having reversed CVC sequences with unattested syllable nuclei, like *[ə u].

18 It was assumed that reversing *individual* segments like /l/ and /n/ would not remarkably alter their perception. The reason for this is that such *continuants* are not characterized by a *sequence of distinct* articulations or *discrete* acoustic units with clear boundaries that listeners need to hear in that *order* to ensure their correct identification of the sounds in question. Instead, as spectrograms show, /l/s and /n/s are *continuous* acoustic events which do not start or change or end abruptly. It was therefore assumed that any misperception of /n/ or /l/ in reversed CVC syllables would be attributable to their unusual position within the *word*, and not to the fact that their acoustic manifestation had been changed remarkably within single reversed *segments*.

19 It was, however, impossible to completely prevent distortion of the stimuli. Reversed words still sounded a bit funny, probably partly because the initial consonants were less strongly articulated than the final consonants, which is the opposite of what one usually expects. Also, one subset of the reversed CVC words (see Table 1, column D) was then played to the subjects also in its original sequence as well (see Table 1, column E; see below for an explanation). Therefore, the words in this set did have a rising intonation.

20 The resulting amplitude was thus the sum, not the average, of the two signals mixed together.

21 Some of these Italian words in *-l* and *-n* end in a consonant by default; others appear in truncated forms, that is, without the final vowel, as is sometimes done in poetry.

22 One important way a consonant's place of articulation can be identified is by means of the formant transitions of its adjoining vowels (see, for example, Stevens and Blumstein 1978). A way to check to what extent identification of [l] and [ʃ] as allophones of /l/ is dependent upon the vowel transitions into and out of it could have been to substitute an [n] or [l] for [ʃ] word-finally or a [ʃ] for a [l] word-initially, while maintaining the original vowel transitions unchanged. This approach, however, was not adopted in this study because of the difficulty of separating vowel nuclei from their flanking consonants.

23 Duration measurements of /l/ articulations in a study by Gartenberg (1984:141-142) reveal that prevocalic /l/ is longer than postvocalic /l/ both in terms of total duration and of onset duration.

24 Actually, some native speakers of English do pronounce [ʃ] word-initially, before high back vowels, as in [ʃ u k] 'look', and children may even be more extreme in this and consistently pronounce [w] instead of [l] in all word-initial positions. (This peculiar phenomenon is called *lall* by speech pathologists; see Tiffany and Carrell 1977:33.) One of my subjects told me that as a kid his mother made(!) him pronounce all word-initial /l/s as [ʃ]s on the ground that she didn't like Spanish-sounding /l/s.

25 In a few cases the transcription provided is more phonemic (broad) than phonetic (narrow). Thus, /t u:/ (column C) and /m u:n/ (column H) are better transcribed as [t u ə ʃ] and [m u ə n], respectively. With regard to the stimuli in column G, the initial [ʃ] of some of them, namely #s 44, 46, and 47, was so long that it sounded like a [u]. Also, except for #s 40 and 48, all the stimuli, when played backwards, sounded as if they had a more central vowel as their syllable nucleus than when played in the original sequence. Thus #s 41 through 47 sounded, respectively, like [w ə n] or [w e n], [w ɒ s] or [w ə s], [w ə ʒ], [w ə ʃ], [w ɒ z], [u ə θ], and [u u: n].

26 In the directions, subjects read that if they got distracted during the experiment, they could ask to have the stimuli they had missed played back to them again at the end of the experiment. One subject missed two stimuli, and another missed three, so those stimuli were played back to them. As in the regular experiment session, they were played three times with a half-second pause before each repetition.

27 Due to a typo that was detected only later in the answer sheet, for item H in the training session, namely [n ʌ l] — reversed from the original sequence [l ʌ n] — one subject had to choose between L and N rather than L and W: She chose L. This mistake was thus instructive, because it showed how a differently-phrased question could deeply influence an answer. That is,

that [t], although definitely more similar to a /w/ than an /l/ — at least according to the other subjects' answers — was still, however, identifiable as an /l/ if compared to an /n/.

²⁸ Cooper, Delattre, Liberman, Borst, and Gerstman (1952:603) mention reasons of possible confusion between /l/ and /w/.

²⁹ The initial segment sounded like [w] or [u], depending on its duration, even to linguistically trained people, even if it was 'only' velarized, and not also labialized, as a [w] would be.

BIBLIOGRAPHY

- Bladon, R.A.W., and Ameen Al-Bamerni. 1976. Coarticulation resistance in English /l/. *Journal of Phonetics* 4:137-150.
- Clark, John, and Colin Yallop. 1990. *An Introduction to Phonetics and Phonology*. Oxford: Basil Blackwell.
- Cooper, F.S., P.C. Delattre, A.M. Liberman, J.M. Borst, and L.J. Gerstman. 1952. Some Experiments on the Perception of Synthetic Speech Sounds. *The Journal of the Acoustical Society of America* 24/6: 597-606. Reprinted in Lehiste, Ilse (ed.) 1967. *Readings in Acoustic Phonetics*. Cambridge, Massachusetts: the M.I.T. Press, 273-282.
- Gartenberg, Robert D. 1984. An electropalatographic investigation of allophonic variation in English /l/ articulations. *Phon Lab Univ Reading Work in Prog* 4:135-157.
- Giles, Stephen B., and Kenneth L. Moll. 1975. Cinefluorographic Study of Selected Allophones of English /l/. *Phonetica* 31:206-227.
- Hardcastle, W. J., and W. J. Barry. 1985. Articulatory and perceptual factors in /l/ vocalisations in English. *Phon Lab Univ Reading Work in Prog* 5:31-44.
- Ladefoged, Peter. 1993. *A Course in Phonetics*. 3rd edition. Hartcourt Brace Jovanovich College Publishers.
- Lehiste, Ilse. 1964. *Acoustical Characteristics of Selected English Consonants*. Bloomington: Indiana University Press.
- Ohala, John. 1986. Consumer's guide to evidence in phonology, ed. by E. Colin and J. Anderson. *Phonology Yearbook* 3:3-26.
- Ohala, John. Forthcoming. Emergent obstruents: diachronic and phonetic data.
- Sproat, Richard, and Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21:291-311.
- Stevens, K. N., and S.E. Blumstein. 1978. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64:1358-1368.
- Tarnóczy, T. 1948. Resonance data concerning nasals, laterals and trills. *Word* 4: 71-77. Reprinted in Lehiste, Ilse (ed.) 1967. *Readings in Acoustic Phonetics*. Cambridge, Massachusetts: The M.I.T. Press.
- Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.