Assimilation of Voice in Dutch

I.H. Slis
Nijmegen, the Netherlands

1. Introduction

In the Dutch literature on the assimilation of voice - which stems for the larger part from before 1960 and is therefore mainly based on non-instrumental investigations - a number of controversies can be observed. Yet, the following generalisations which are relevant for assimilation of voice can be made, be it with some reservation:

a. As a rule, syllable-final obstruents are voiceless in Dutch. Therefore, assimilation of voice can only occur if a voiceless syllable-final obstruent is followed by a voiced syllable-initial consonant.
b. If the second consonant of a two-consonant cluster (= C2) is a voiced fricative, progressive assimilation can nearly always be observed.
c. If C2 is a voiced plosive, regressive assimilation can be observed in the majority of cases.

Relatively speaking, the third rule is the most controversial one: also, progressive assimilation is reported frequently in clusters with a voiced plosive as C2. A number of causes are suggested for this devoicing, such as the emotional state and sex of the speaker, stress position, informational load of the following word and the phonological composition of the cluster.

Our major objective in the present study is to survey various influences on assimilation of voice. To this end, we make a distinction between:

a. non-linguistic influences such as organic conditions of the vocal cords, and
b. linguistic influences, such as the phonological or stress context.

1.1. Non-linguistic influences

At least some of these factors concern the way in which the vocal cords vibrate. In this paper we will pay attention to the influence of vocal pitch, voice quality and sex of the speaker.

Voice activity depends on conditions within the larynx. If, for example, the vocal cords are too far apart or too tense no voicing will occur. During the production of a consonant cluster the airflow through the oral passage can be obstructed to such an extent that the boundary conditions for vocal fold

vibrations are approached. We presume that in a number of cases raising the pitch by increasing the tension of the vocal cords leads to a situation in which these boundary conditions are surpassed, and consequently to an increase in the number of interruptions of voicing. If these interruptions start within 50 ms after the preceding vowel and continue until the moment of oral disclosure (see operational definition below), progressive assimilation takes place. Another factor that can be expected to influence boundary conditions of vocal cord vibration is the organic state of the vocal apparatus. It is very likely that in speakers with a poor voice the organic state is less favorable for vocal cord vibration than in speakers with a good voice.

In addition we expect differences in assimilation between the speech production of men and women, since male and female larynges are known to be of different size.

Another factor which might influence the degree of assimilation is speech rate. Successive speech gestures will overlap to a greater extent with increasing speech rate, resulting in a higher proportion of assimilated clusters.

1.2. Linguistic influences

Less obvious is the relation between voicing and assimilation on the one hand and linguistic variables on the other. In this paper we will study the influences of place and manner of articulation, linguistic stress and the phonological length of the preceding vowel.

From a comparison of the voice onset times of dental and labial plosives it appears that there is a relationship between place of articulation and voicing. In almost all languages studied voice onsets in dentals are delayed compared with those in labials (Lisker and Abramson, 1964).

As far as manner of articulation is concerned, differences with respect to voicing are to be expected between clusters containing a fricative and a stop and clusters containing only stops. By means of transillumination of the glottis we observed a larger glottal aperture during the production of fricatives than during that of stops (Slis, 1970). This may have consequences for the assimilation of voice, because the larger glottal opening may lead to a higher proportion of interruptions of voice activity.

Since stress seems to be effected by a greater articulatory effort (Slis, 1971), we expect that the voice character of a consonant belonging to a stressed syllable dominates one belonging to an unstressed one. This implies that in pre-stress condition, in which the voiced C2 belongs to the stressed syllable, regressive assimilation is most probable. On the other hand we expect progressive assimilation to occur when the voiceless C1 belongs to the stressed syllable.

Furthermore, voiceless consonants tend to be longer and are pronounced with more effort after short than after long vowels. Consequently, we expect a greater prominence of the voiceless element in clusters after short than after long vowels which will lead to more progressive assimilation after short

vowels. In view of the above relations we tentatively suggest that assimilation of voice is not a mentally programmed process, but rather the result of unintentional articulation.

2. Method

2.1. Independent variables

We investigated the following variables which are directly related to voice activity or coarticulation:

1. Pitch: normal intonation vs. low, medium and high monotonous speech.
2. Voice quality of the speakers: good vs. poor quality.
3. Sex of the speakers: male vs. female speech.
4. Speech rate: slow, normal and fast speech.

In addition, we studied some other variables whose relationship with voice activity is less obvious, but which may contribute as argued above:

5. Phonological composition of the clusters with respect to:
   a. manner of articulation: stop-stop vs. fricative-stop.
   b. place of articulation: /p,t,k,f,s,χ/ followed by /b/ or /d/.
6. Stress conditions: before or after stressed syllable or without stress on the adjacent syllables.
7. Phonological ength of the preceding vowel: /a:,o:,ø:,e:/ vs. /ɑ,ɔ,œ,I/.

2.2. Speech material

The production of words can be regarded as the performance of sets of coherent motor commands. In this light, assimilation within words may be regarded as engrained motor patterns. This is not the kind of assimilation we want to study in this paper. We will restrict ourselves to assimilation across word boundaries, where a 'new program' is needed since each combination is unique. Clusters in which C2 is a fricative are also excluded from this report. Since nearly 100% of these clusters show progressive assimilation (Slis, 1981) we do not expect that the few deviations from the rule will significantly contribute to our experiment.

For these reasons we only performed experiments on fricative-stop and stop-stop clusters across word boundaries, the second (stop) consonant being voiced. Seven different series of measurements were carried out on clusters in sentences that were read aloud. In each of the series between 20 and 40 sentences were recorded of between 20 and 45 subjects. In this way we obtained a total of 4661 two-consonant cluster-tokens.

2.3. Measurements

From preliminary observations we knew that assimilation, either regressive

or progressive, need not always be complete: a large number of intermediate cases occurred. We therefore needed a strict definition of regressive and progressive assimilation. This definition should be based on the general definition which states that we speak of assimilation if one sound segment exerts influence upon the articulation of another, so that the two sounds become more alike or even identical (Crystal, 1980:35).

From U.V.-oscillograms (paper speed 50 or 100 mm/s) of the speech signal we derived the moments of closing and opening of the vocal tract from the moments of sudden changes in the amplitude. From electrolaryngograms we deduced the moments of opening and closing of the glottis; in nearly all cases these moments occurred at the end and the beginning of observable voice activity. In few cases the amplitude of the larynx signal was not completely reduced to zero. We assumed that the low-amplitude part of the electrolaryngogram depicted edge vibrations of the vocal cords with an open glottis. In these cases we made a 'best guess' based on the slope of the amplitude envelope.

From previous measurements on intervocalic voiceless plosives (Slis, 1970) we learned that voicing may continue about 20 ms after the moment of closing the oral tract. We called this continuation of voice a 'voice tail'. The moment of voice onset (VOT = voice onset time) occurred at or after the moment of opening of the vocal tract. We defined a syllable-final obstruent as voiceless when the voice tail was shorter than 50 ms, and an initial consonant as voiceless when the VOT took place at or after the moment of opening the vocal tract. Although these moments could not be determined with a high degree of precision (glottal period duration 5 to 10 ms) only a few cases proved to be ambiguous.

To sum up we arrived at the following definitions of assimilation categories:

1. No assimilation: The voice tail is less than 50 ms and VOT is negative.
2. Regressive assimilation: Voicing continues during the closed interval or the voice tail is longer than 50 ms. (interruption of voicing after so long a voice tail is attributed to aerodynamic causes and not to glottal opening).
3. Progressive assimilation: The voice tail is less than 50 ms and the VOT is zero or positive.

3. Results

3.1. Pitch

In line with the expectation mentioned above the results showed an increase of progressive assimilation with increasing pitch from 20% to 32% ($\chi^2 = 24$, $df = 4$, $p < .001$). Additionally, more progressive assimilation was found in low-pitched monotonous speech (20%) than in normally intonated speech (10%). We assume that this is brought about by a higher tension of the laryngeal musculature in monotonous speech in order to keep the larynx in a fixed position.

3.2. Good vs. poor voices

The results of our experiments, based on 1103 clusters, showed about twice as much progressive assimilation in poor (27%) as in good voices (13%, $\chi^2 = 44$, df = 2, $p < .001$). This result confirms our expectation on this point.

3.3. Male vs. female voices

The differences in assimilation between the speech of men and women were studied in four experiments. In total 708 clusters of female speakers were obtained. These were compared with the pooled results of the male speakers in all experiments. In three of the four experiments, the results showed that women have significantly more progressive assimilation (45%) than men (25%) in all the conditions examined ($\chi^2 = 140$, df = 2, $p < .001$).

3.4. Speech rate

In one experiment (973 clusters) we measured the influence of speech rate in clusters in a pre-stress position comparing slow, normal and fast speech. With all cluster types studied we found a significant increase in degree of assimilation going from slow, via normal, to fast speech ($\chi^2 = 122$, df = 4. $p < .001$). This increase was from 51% to 92% when C2 was a stop and from 77% to 100% when C2 was a fricative. Normal speech took an intermediate position. No remarkable change in direction of assimilation could be observed.

3.5.a. Phonological composition with respect to manner of articulation

A comparison of stop-stop with fricative-stop clusters for the summed data of all experiments shows highly significant differences in assimilation ($\chi^2 = 162$, df = 2, $p < .001$) in all the conditions investigated, i.e. before, after and without stress, in the speech of both men and women. More assimilation was found in stop-stop clusters (83%) than in fricative-stop clusters (77%). In fricative-stop clusters we observed more progressive assimilation (40%) than in stop-stop clusters (30%).

3.5.b. Phonological composition with respect to place of articulation

Clusters in which C2 was a /d/ showed more progressive assimilation (38%) than those with a /b/ (32%) in all the conditions investigated (in which manner of articulation of C1, the stress condition and the sex of the speakers, were variables). Five out of eleven differences were significant ($\chi^2 = 60$, df = 2, $p < .05$).

3.6. Stress condition

In all the conditions investigated (4236 clusters) we observed highly significant differences in assimilation ($\chi^2 = 276$, df = 2, $p < .001$): before stress we found twice as much regressive assimilation (54%, vs 24% progressive assimilation). Conversely, we found that after stress the frequency of progressive assimilation (50%) was twice that of regressive assimilation (19%). This confirms our expectation.

3.7. Phonological length of the preceding vowel

In one of the experiments we investigated the influence of the length of the preceding vowel on the assimilation of a cluster. Although we did not find significant differences, there was a tendency towards a higher frequency of assimilation after long stressed (65%) than after short stressed vowels (56%, $\chi^2 = 5.5$, df = 2, $p < .10$). This was mainly attributable to the higher frequency of regressive assimilation (26% after long and 13% after short vowels). This result is in line with our prediction mentioned above.

4. Discussion and conclusion

All the expectations put forward in the introduction are confirmed by the data. These expectations were based on the assumption that assimilation of voice is a result of unintentional coarticulation rather than a mentally programmed process. Alternative explanations for our results are possible. However, they are heterogeneous; for one aspect there is one alternative, for another we have to present another alternative. For instance, with respect to the influence of pitch, the only plausible explanation is the organic state of the vocal cords as put forward in the introduction. Differences in assimilation in the speech of men and women, however, may also be a sociolinguistic phenomenon. similar to other sex differences in language (e.g. Smith, 1979). Differences due to linguistic influences may also be the result of the intentional application of different rules depending on the composition of the cluster, or its phonological context.

Although no convincing proof can be given that all the effects which we observed are consequences of coarticulation, we feel attracted to an esplanation that covers all oppositions studied; coarticulation seems to be the only possibility.

References

Crystal, D. (1980). *A first dictionary of linguistics and phenetics.* London: André Deutsch.

Lisker, L. and Abramson, A.S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word,* **20**, 384-422.

Slis, I.H. (1970). Articulatory measurements on voiced, voiceless and nasal consonants: A test of a model. *Phonetica,* **21**, 193-210.

Slis, I.H. (1971). Articulatory effort and its durational and electromyographic correlates. *Phonetica,* **23**, 171-188.

Slis, I.H. (1981). The effect of speaking rate on assimilation of voice. *Proceedings of the Institute of Phonetics Nijmegen,* **5**, 160-176.

Smith, P.M. (1979). Sex markers in speech. In: K.R. Scherer and E. Giles (eds.) *Social markers in speech.* Cambridge: Cambridge University Press, 109-146.