UNIVERSITEIT VAN AMSTERDAM

Research Master's thesis
Academic year 2011/2012

Diachronic change in the postvocalic /r/ in the Dutch of Amsterdam.

A synchronic approach to detect an ongoing sound change.

Dhr. Prof. dr. P.P.G. Boersma
Faculty of Humanities - Chair group Linguistics

Dhr. dr. D.J.M. Weenink
Faculty of Humanities - Chair group Linguistics

Daniele Dalmasso
6099734

1. Introduction

Several unalike realisations of /r/, especially in postvocalic position, are commonly heard in Present-Day Dutch. This case of allophony has not been overlooked by scholars. A number of studies on the variability in the production of the phoneme /r/ are found in the literature. According to Booij (1995), /r/ can be realised in at least six different ways as the consequence of 'individual and regional variation' (Booij 1995:8). Van de Velde (1996) lists ten variants of the r-sound in postvocalic Dutch. It is considered 'an obscure sound in the sense that it is subject to almost extreme inter- and intra-speaker variation' (Smakman 2006:222). Apparently, 'the most variable sound in the Dutch language area is probably /r/ but, due to its complexity, is also one of the least studied' (van de Velde, van Hout et al. 1997:369).

Most of the literature on this issue is mainly focused on the description of the different phonetic realisations (providing lists of sounds) and on their distributions across the Netherlands. Apparently, the only work that tackles the /r/ allophony from a diachronic point of view is provided by van Bezooijen et al. (2002) which tested the different pronunciations of postvocalic /r/ produced by two groups of speakers from the cities of Haarlem and Nijmegen. Significant differences have been found for the frequencies of use of approximant sounds between speakers of the two cities and between young and old speakers.

Likewise, this research aims to discover patterns of variation about this interesting sound among the speakers of the Dutch spoken in Amsterdam (from now on D.o.A.). Our hypothesis is grounded, both on these studies and on the impressionistic evaluation about a change in progress that seems to bring a more stable sound for the V/r/ phonological slot. That is, from a broader set of realisations in which the rhotic trill is probably the most common sound, an underway sound change conveyed by young adults is leading to a more fixed realisation, usually a rhotic approximant.

To test whether the hypothesis of the ongoing change is true, an apparent-time based study has been carried out: a synchronic approach to language to catch a diachronic change.

Section § 2. is about some important definitions for the rest of the work. Section § 3. gives an outline of rhotic sounds whereas section § 4. sketches a portrait of the /r/ sounds in the modern history of Dutch. Section § 5. regards language change and introduces the reader to the apparent-time study whose application is fully set out in section § 6.. The methodology applied is provided in section § 7.. The procedures of the analysis, preliminaries and cues for the acoustic classification, are in section § 8. and continue in section § 9. where all the rest of the analysis, by eye and by computer, is described. Section § 10. shows the results and their statistical significance. Discussion and comments on the outcomes are provided in section § 11.. Conclusions are presented in section § 12..

2. Definitions

The present investigation aims to verify whether a sound change is taking place in the D.o.A. In order to make things clearer, this section explains several concepts applied throughout this work.

First of all, since this thesis is on sound change, it is essential to go through its definition, that is 'any appearance of a new phenomenon in the phonetic/phonological structure of a language' (Lass 1984:315). Therefore, any modification on the way a sound is produced by the phonatory apparatus is considered a sound change.

Citing both phonetics and phonology, the above definition approaches the phenomenon of change from a quite wide position. As far as this work is concerned, the affected level on the supposed change of the /r/ in Amsterdam is the phonetic one. The difference between the two subjects is too important to be overlooked.

In the phonetic realm the change concerns just the 'physics' of the sound: that is, a change on the production of the sound through a different pronunciation e.g. duration of the sound, modification of the position of the tongue or the like. In this case, there are not consequences for the sound inventory of the language. The phonological level, in turn, deals with the abstract representations of the sounds stored in the minds of the speakers. A phonological (phonemic) change, triggered by a sound change, would modify the number of phonemes in the inventory (e.g. increased in cases of splits, or reduced in cases of mergers). Since the different allophones of /r/ described here all fall into the category of the phoneme /r/, that is the mental representation of the Dutch <r>, we deal with a phonetic (non-phonemic or allophonic) change.

Another important dichotomy splits sound change phenomena into conditioned and unconditioned changes. Since the supposed sound change does not affect /r/ in all contexts, this research deals with a conditioned change: the pronunciation of /r/ clearly shows a great deal of variability only in coda/postvocalic position. The coda is, 'the portion of a syllable which may follow the syllabic nucleus, e.g. the /p/ of /kʌp/ 'cup" (Crystal 2008:82). Seeing that in Dutch more than one consonant can follow the vowel (nucleus), the position aimed to is always the postvocalic one. The environments investigated here are therefore both V/r/# and V/r/C#.

A couple of further definitions that must be made clear are the following: */r/ sound* in this work can mean both a rhotic sound or any sound that fills the phonological slot of the phoneme /r/. The term *rolled sound* is meant to gather both trills and tap sounds.

3. An outline of rhotic sounds

Not only rhotics sounds are recorded among the possible realisations in postvocalic position in present-day Dutch but also fricatives and the schwa (see table 4.1 in section § 4.). However, rhotics are possibly the most likely sound to occur in this position which phonologically is represented by /r/. However, what is really a rhotic sound? The class of sounds referred as r-sounds or rhotics is under debate among specialists. According to Ladefoged and Maddieson (1996) at least eight different realisations fall into this category (table 3.1). What is still unclear for phoneticians is what actually makes these sounds a unitary group that stands out from the other sounds. The conclusion drawn by the extensive study of Ladefoged and Maddieson is rather pessimistic: '[a]lthough there are several well-defined subsets of sounds (trill, flaps, etc.) that are included in the rhotics class, the overall unity of the group seems to rest mostly on the historical connections between these subgroups, and on the choice of the letter 'r' to represent them all' (Ladefoged and Maddieson 1996:245).

Another widely cited statement on this matter says that: 'There is no physical property that constitutes the essence of all rhotics' (Lindau 1985:166).

Definition	IPA symbol
Voiced dental or alveolar trill	r
Voiced dental or alveolar tap or flap	ɾ
Voiced dental or alveolar approximant	ɹ
Voiced post-alveolar flap	ɽ
Voiced post-alveolar approximant	ɻ
Voiced uvular trill	ʀ
Voiced uvular approximant	ʁ
Voiced dental or alveolar lateral flap	ɺ

Table 3.1. r-sounds overview. From Ladefoged & Madison (1996).

The lack of a unique property that satisfactorily designates this class and makes it stand out from the rest of the other sounds is a limitation that will reflect on the problems for the acoustic classification of this work (see sections § 8.2., § 9.3 and § 9.4.).

Phonologically, however, these sounds seem to behave in similar ways across languages. An estimated 75% of the world's languages have at least an /r/ sound in their phonological systems (Maddieson 1984). One fifth of this group have two or more variants (van Bezooijen, Kroezen et al. 2002:1). Dutch falls under this group, showing quite a range of allophony in the rhotic speakers' repertoires.

4. Postvocalic /r/ in Dutch, an outline

A short introduction to the history of /r/ in Dutch is worthwhile being mentioned. From the descriptions on the Dutch language of the 19th century, there seems to be little space for doubting on the trilled realisation of /r/ although the potential effects brought by the position of the sound in the syllable was generally ignored (Smakman 2006). The following century saw in the first place the emergence of variants of /r/ realised in the back of the mouth, shifting thus the place of articulation from alveolar to uvular (without changing the manner of articulation though) through a stunning influence of French. Towards mid-century, this variant turned to be widely accepted in the language posing 'increasing uncertainty regarding the proper standard pronunciation of (r)'(Smakman 2006:221). Nowadays, a relevant number of allophones of /r/ in postvocalic position are heard. Among these, an approximant realisation seems to spread and gain acceptance by the speakers (Booij 1995; van den Heuvel and Cucchiarini 2001; van Bezooijen, Kroezen et al. 2002; Smakman 2006; Scobbie and Sebregts 2010). The change, that allegedly first surfaced in the *'t Gooi* area in the North Holland province and it seems to gain acceptance among the speakers across the country through the influence of the media, which is mostly set in this area of the Netherlands (van Bezooijen, Kroezen et al. 2002; Smakman 2006).

The picture however, is far from clear. In fact, even more realisations have been recorded throughout the accents of Dutch making this sound so interesting because its astonishing number of allophones: 'almost all variants of /r/ found in the languages of the world are observed in the Dutch Language area' (Verstraeten and Van de Velde 2001:45).

In his 1996 survey Van de Velde detected no less than ten different realisations of /r/.

Definition	**IPA symbol**
Voiced dental or alveolar trill	r
Voiced uvular trill	ʀ
zero	Ø
schwa	ə
retroflex sounds	ɽ ɻ
tap	ɾ
Voiced dental or alveolar approximant	ɹ
Voiced uvular approximant	ʁ̞
Palatal approximant	j
velar sounds	x χ ʁ

Table 4.1. The ten Dutch variants for V/r/ according to van de Velde.

5. Language change (get over it)!

It is nothing surprising that people talk differently. Here, however, we are not concerned with distant languages spoken in two remote corners of the Earth. Instead, this study deals with the same national language spoken by a group of speakers selected in the very same city. Many potential variables are at play when speech is produced by the speakers of a community. Such can be sex, race, social status, religious affiliation and among others, the one we think is responsible for the change we are investigating, that is age.

It is not unusual that people often notice or even criticise that their own language is changing. Apparently, speakers have an innate negative attitude toward language change. Think of the comments of the Roman writers on the corruption of Latin or looking at the present at the *Academie Française*, which among other duties, settles whether a new word can enter into the French language or not according to a strict conservationist policy. (More on attitude towards change is available, among others, in (McColl Millar 2007) and (Aitchinson 1991)). Intriguingly enough, also the case of the rise of the approximant /r/ in Dutch led someone to complain about it. Possibly the oldest reference to the phenomenon is found in Meertens (1938) who criticised it as the "unrecognisable and distorted' /r/ of some radio presenters' (van Bezooijen, Kroezen et al. 2002:2).

Yet, we should not be surprised at all if we all agree that '[t]ime changes all things: there is no reason why language should escape this law' (Saussure de 1915/1959:77). The fact that individuals may able to spot that something in the language is 'changing', 'not longer said as before' will be extremely important for the following discussion: can we see language change while it is taking place?

5.1. Is catching a change in progress a feasible task?

McColl Millar (2007), introduces the field of Historical Linguistics by looking at the changes that led Old English, a language spoken some eleven centuries ago, to turn into Present-day English. However, he takes as a first assumption that 'the people who spoke this language [Old English] taught it to their children, who *taught* it to their children, who TAUGHT it to their children, who… until it finally reached us. But it reached us in a very different state' [his emphasis](McColl Millar 2007:2). The cycle of the language that flows through generations is indeed what daily occurs since the appearance of language among human beings. Yet, languages indisputably change despite the assumption that, in normal conditions, parents are perfect 'teachers' and children are perfect 'learners'.

Literary traditions give us clear evidence for the unstoppable force of language change. Taking the aforementioned case of the English language for instance, it will be enough to compare the

Anglo-Saxon Chronicle, written in Old English, and the language employed by Chaucer some three centuries later (Middle English), to observe huge discrepancies. In turn, the language of the *Canterbury Tales* is astonishingly different from the one in which the latest *Harry Potter*'s novel was recently issued (Present-Day English). Hence, assuming that the 'teaching-learning-teaching' cycle is immune from deviances and observing the obvious fact that languages change we face a major problem: change seems unobservable. Apparently, therefore, the work of the historical linguist is limited both by the course of time and by the availability of historical records. Language change would be therefore observable only when linguistic discrepancies emerge from written records spaced by decades or centuries.

Yet, at some point an innovation must be initiated by someone in a community, then it has to spread from speaker to speaker. After the initiation thus, the change is implemented by the speech community. Given that, one may wonder whether language change is observable while it takes place. If our premises are true, we have to assume that from a linguistic stage Time1 to a linguistic stage Time2 a period of 'development' must take place. This is the reasoning behind the efforts of modern linguists that tried to challenge the diachronic-synchronic dichotomy asserted by de Saussure 'the opposition between the two viewpoints… is absolute and allows no compromise. A few facts show what the difference is and why it is irreducible' (Saussure de 1915/1959:83).

The period spanning between T1 and T2 was first named by Weinreich et al. (1968) as the *transition problem* which was raised along with other 'problems' concerning how and why a change may occur and spread in a language. Having said that, the questions are: can we detect an ongoing change in the so called period of transition, that is, in a single point in time? In other words, is it possible to conduct an investigation that ascertains a change by analysing synchronic data?

We have already seen that de Saussure was sure that it was not. Hockett (1958) and other linguists with him, contended that 'the actual process of language change is unobservable – it can only be detected through its results' (Bailey 2002:312). 'Only indirect methods could show us sound change in progress' (Hockett 1958:444).

5.2. Apparent-time method as the solution to the 'transition' problem

The nuisance brought by the invisible hand of change was not left unchallenged and new techniques have been sought and applied to eventually overcome its supposed undetectability: 'sociolinguistic studies have now convincingly shown that sound change in progress *can* be observed' [her emphasis] (McMahon 1994:49). A turning point therefore is emerging here: historical linguistics leaves room to sociolinguistics '[w]hile historical linguists have traditionally studied language change as a diachronic succession of completed changes reflected in the historical records, sociolinguists have preferred to study change as an active process reflected synchronically in age-based linguistic variation' (Boberg 2004:250).

The original approach is therefore unaltered, language variation investigation must necessarily be comparative, but it is the perspective that changes through a different collection of the evidence: from the 'armchair method' to a 'tape-recorder method'. In other words, evidence for change is no longer sought in libraries but in the real speech of the world through a microphone.

We already revealed that the breakthrough came out of sociolinguistic studies, but how? In order to detect linguistic idiosyncrasy data should be drawn from samples of the same population of speakers at two different points in time. One way to deal with this would be repeating the same interview to the same informants after a period of time, say, ten or twenty years. The application of this method, is barely feasible. In *real-time* studies one has to return in the same community after a considerable number of years to interview those speakers already interviewed at the time of the first visit (panel study) or a second diverse but representative sample of speakers (trend study). It goes without saying that both types of studies in real-time are very difficult to apply: results are available only after several years of waiting, hardly any institute of research would dispose funds for such a type of slow investigation, etc. Moreover, it would not be correct to call it a synchronic approach since a period of time must span between the two time points.

Another approach to language change, totally synchronic, has been successfully developed and applied in the works of William Labov in the 1960's. The main breakthrough was the idea that when social and stylistic factors were held constant, any linguistic variance between generations is the signal that an ongoing change is taking place. In doing so, he achieved in proving that change is indeed observable in its transition period also synchronically through a rather economic research compared with real-time studies. The *apparent-time* construct assumes that 'a change in progress is under way and that the variant most characteristic of older speakers' speech represents the earlier stage and the variant more typical of younger speakers' speech shows what it is changing to' (Campbell 2004:220). The assumption underlying apparent time is that 'most people learn their language in childhood, and that, after adolescence, they do not normally introduce any further significant changes into their speech' (McColl Millar 2007:362) and 'unless there is evidence to the contrary, differences among generations of adults mirror actual diachronic developments in a language when other factors, such as social class, are held constant' (Bailey, Wikle et al. 1992). The answers are thus provided by the application of this brilliant technique: if the young speak differently from the old, other things being equal, then age is responsible for variation.

Several studies of this type have been conducted after the earliest Labov's *Martha's Vineyard* and *New York studies*. These two researches must necessarily be mentioned because they traced the line to follow for sociolinguistic based studies. Based on McColl Millar's discussion (2007:342-343), I put forward five points that justify the reason why Labov's works have been so important and influential:

- They introduced a new era, they are the breakthrough of sociolinguistic studies
- They defeated the skeptical position of those linguists who dismissed the transition problem as impossible to overcome
- They introduced the use of statistical analysis
- They developed the notion of social stratification, 'with members of different groups using different values of the variable' (McColl Millar 2007:343)
- They showed that variation means change in progress.

5.3. Some basic assumptions on change

'Changes typically begin with variation, with alternative ways of saying the same thing entering the language' (Campbell 2004:219). Sociolinguistic researches are thus focused on how and why a particular change arises in a language.

Usually, changes are already at the disposal of the speakers and are not 'comparable to meteorites falling from the sky. They usually originate from elements already in the language which get borrowed and exaggerated' (Aitchinson 1991:74).

According to general claims about linguistic change based on large-scale investigations 'women lead most linguistic changes (women accept and help to propagate the linguistic changes earlier than men do)' (Campbell 2004:221). If the change is above the level of consciousness these 'are usually in the direction of speech forms with overt prestige' (Aitchinson 1991:75).

All these points will be discussed throughout the work: we will see whether the supposed change of the postvocalic /r/ in the D.o.A. follows the most common assumptions about change.

6. Application of the apparent-time model for this study

6.1. Detecting the variable

The variable under investigation is the articulation of the phoneme /r/ in postvocalic position in the D.o.A. The assumption (see § 4.) is that the sound is produced by old speakers more likely as a trill [r], [ʀ] whereas young speakers there is a strong tendency to pronounce it as a rhotic approximant sound [ɹ] [ɻ]. The old variable is then the trill-like pronunciation whereas the young one is represented by a rhotic approximant-like sound: the apparent-time construct will shed light on the progress of the change. This variation is possibly the best issue that the Dutch language is offering for a work such as the apparent-time model.

The work will thus be addressed exclusively on the phonetics of the sounds collected through a test administered to a group of native speakers of Dutch. This test, as it will be shown in the following sections, takes into account all possible phonological contexts in which the postvocalic /r/ may fit in order to see whether different conditions (such as presence or absence of stress in the previous syllable, quantity and quality of the previous vowel) may affect the phonetic production of the sound.

Along with our main target, that is the /r/s, another phonetic phenomenon has gained our attention and will be not overlooked. That is the insertion of [ə] in the V/r/C# contexts when C is a labial or velar sound (see sections § 9.3.1. and § 10.2.6.). This a typical phenomenon in Dutch and it is important for two reasons. Firstly, it invalidates the phonological context we are aiming to, because it makes it intervocalic rather than postvocalic, so each occurrence of this type will be discarded from the main analysis. Secondly, it is interested to discern whether the epenthetic schwa is correlated (positively or negatively) with the change /r/ [trilled] > /r/ [approximant].

6.2. Selecting the informants

Once the linguistic variable has been identified, the following step in the application of this technique concerns the selection of the informants. According to that very important quotation by Bailey et al. (1992) in section § 5.2., we need to pay a great deal of attention to the participants of our test so that 'other factors, […], are held constant'. It is one of the most important assumptions for obtaining reliable results from an unbiased investigation based on the hypothesis that age is responsible for change. The 'choice' of the informants is thus important as much as difficult. In fact, if we want to ascribe to age the difference in pronunciation between two groups of speakers, we need to fix to the same levels all the other variables i.e. gender, educational background, place of birth, social status, and the like.

Only in so doing we can safely address age as the factor for the change. Otherwise we would be forced to tempt to determine other correlations due to other variables: change because of social status, change because of ethnicities, and so on. Indeed, this task is too complex to be accomplished thoroughly without pitfalls. The informants from which this work relies on is given by speakers with the following characteristics:

- Both males and females have been sampled.
- The young group spans between 18 and 27 years old, whereas the old one between 51 and 70.
- Most of the subjects have received a high education.
- All of them are native Dutch speakers with Dutch parents[1].
- Most of them were born, raised and still living in Amsterdam. Some of them grew up somewhere else in the Netherlands but moved to Amsterdam only after their childhood. In

1 Two participants were bilingual because one of their parents was not Dutch.

turn, some others grew up in Amsterdam but spent a remarkable number of years outside. These are indeed potential limitations (see section § 10.2.5.).

7. Methodology

7.1. The target sound: the phonological clusters

Given that this work focuses on the realisation of /r/ in coda-postvocalic position it is essential to study whether the different environmental clusters may play a role on the realisation of the r-sound. The starting point is then represented by the study of the vocalic inventory of Dutch. According to Booij (1995), the phonological system of vowels in Dutch is made up by 7 long vowels, 5 short vowels and the schwa.

Typology	**IPA symbols**
Short (V)	ɪ, ɛ, ɔ, ʏ, ɑ
Long (VV)	i, y, u, e, ø, o, a
Schwa	ə

Table 7.1. Vocalic inventory of Dutch. Booij (1995).

This distinction however, is not set entirely on a phonetic ground but, rather, to a more phonological one. Close vowels (i, y, u), in fact, are reported to be as long as short ones in Dutch, showing a duration of about 100 ms (Nooteboom 1972:45-7). Nonetheless, they are lengthened ‘before /r/ the high vowels do have the same (extra) length as the other vowels’. This notation is crucial for this research enabling us to keep the close vowels in the long vowel category.

Other two variables that may affect the pronunciation of the r-sound are both the presence or absence of the stress on the preceding vowel and the position of the /r/ within the syllable. Table 7.2 collects the three variables that will shape the origin of the speech material.

Vowel length	Stress	Final context
Simple	+	V/r/#
Long	-	V/r/C#

Table 7.2. The three variables that shapes the reading material.

The consonants in final position can be unvoiced stops /p, t, k/, unvoiced fricative or affricate / s, f, x, ʦ/ or nasals /m, n/. We can now draw a typology of the clusters investigated on this work through the combination of the three variables.

Stressed	Unstressed	Stressed	Unstressed
ˈVr	˘Vr	ˈVrC	˘VrC
ˈVVr	˘VVr	ˈVVrC	˘VVrC
	ər		ərC

Table 7.3. Typology of the clusters investigated.

Considering the Dutch vocalic inventory and the number of different environments, theoretically the contexts to be analysed should be 28 for long vowels (7 [see table 7.1]*2[stressed or unstressed]*2[in final position or not]), 20 for short vowels, (5*2*2) and 2 for /ə/ (1*1*2)([ə] is always unstressed). However, some of them seem impossible in Dutch and no examples were found[2]. Additionally, some others do exist but are very rare.

Applying the typology of the clusters shown in table 7.3. to the possible phonological environments in present-day-Dutch, the experiment consists of 24 environments for long vowels, 16 for short vowels and two for the schwa resulting in 42 different phonological contexts. For each context five words were randomly selected from the dictionary giving, in total, 210 words.

7.2. Speech material

These words are the core of the experiment and they embody the two tasks elaborated for the experiment: a word list and a cueing screen elicitation task.

Out of the five words for each environment, four of them enter into the word list whereas only one, possibly a noun, is singled out and inserted in a cueing screen task. In the environments /yr/ and /ør/ there were no words suitable for being elicited through an image, therefore all five words ended up in the word list.

Dispersed among the target words in the word list, 103 distracter words were added. The choice of these words was completely random unless for the conscious decision of avoiding words containing /r/ in final position. The word list thus presented a total of 171 target words which were scrambled with the distracters for a total of 234 words. Given its length, the list was divided into four sections to allow speakers to take a short break between them.

The second task was an elicitation task: the informants had to speak of what they saw on a laptop screen. The task was made up of a 39-slide slideshow presentation. Each picture represented a target word. Underneath each picture, one or more letters of the target word were given as a hint for

2 Not found: /˘irk/, /˘yrC/, /˘urC/, /˘erC/, /ˈʏr/, /˘ɪr/, /˘ʏr/, /˘ɪrC/; rare: /ˈɪr/, /ˈɪrC/.

the informant. Additionally, it was asked that participants embed the word in a small sentence such as: *het is een*… 'this is a …' or *hier zie ik een*... 'here, I see a…'.

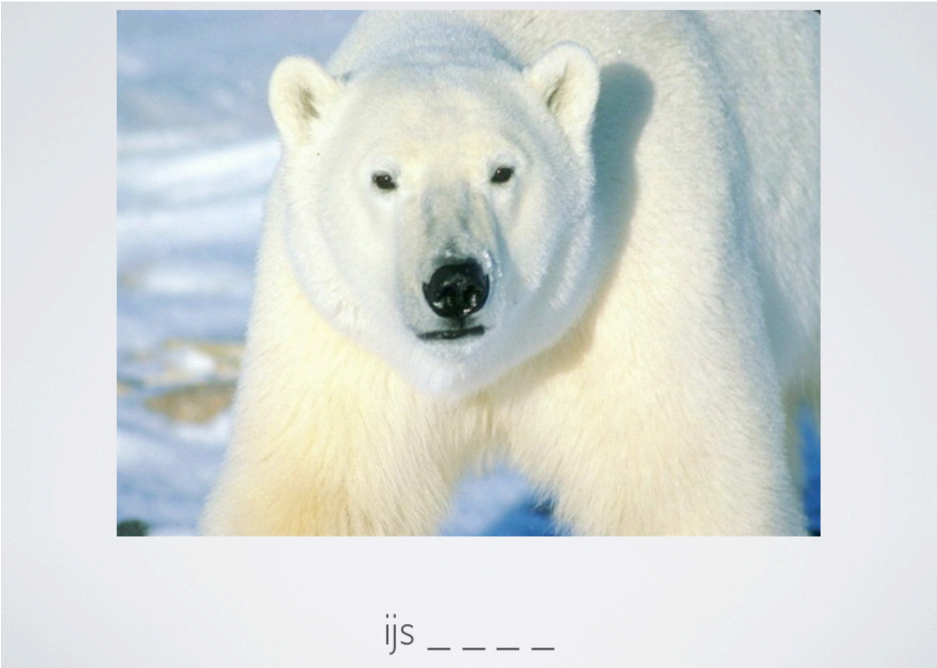

Fig. 7.4. A picture taken from the cueing screen elicitation task. Part of the beginning of the words was given as a hint for the participants. In this case the expected reaction was: "*het is een ijsbeer*" 'this is a polar bear'.

7.3. Consideration on the types of elicitation task

In recording informants' speech, it is important that the researcher gets the actual speech of the informant. To put it simply, in an interview setting it could occur that the informant feels uncomfortable and stops behaving naturally. They may therefore tend to adjust her speech to a more 'formal', 'correct' or 'standard' form. This is exactly what the researcher wants to avoid because what really matters is the real, everyday, casual speech of the speaker.

The problem above stated is well defined in the Observer's Paradox: 'sociolinguists would like to know how speakers speak when they are not being observed, but the only way to find out is by observing them' (McMahon 1994:234).

Different techniques have been conceived to overcome the risks brought by the Observer 's Paradox. Casual speech for instance (talking freely about an informal subject) is one of the safest ways to deal with the risk of getting 'unnatural' speech. Unfortunately, this could not be done because of at least two relevant problems. First, in natural speech we may take the risk of getting too few target sounds and second, the limited knowledge of Dutch by the researcher would have obviously brought problems to the communication stream.

The choice of the word list was therefore a necessity basically because the casual conversation in Dutch could not be covered. This technique, that of a forced reading, may hide some attempts at unnatural pronunciations by the informants. However, the insertion of many distracters prevents the speakers of spotting the purpose of the test, lowering this type of risk.

The cueing screen conversely is free from distracters thus each picture represents a targeted word. The lack of distracters, however, seemed not to affect the pronunciation of the speakers. This type of elicitation task, is in fact less boring than the previous one and informants were interested in finding the word that matched the picture. This sort of challenge made speakers clearly more interested in the task itself rather than in their own pronunciation (although they were aware that it was a recording session focused on their speech).

Thus from the second task, I believe that any risks of getting unnatural samples are very low. Its downside however was obviously the narrower choice of target words since many words are not nouns (so the lower number of items, 39), and indeed we cannot administer the participant a picture of, say, an adjective.

7.4. The test

The recordings were made with a solid state recorder *Marantz* PMD660. Forty speakers (13 girls, 10 boys, 9 women and 8 men) were asked to read the word list closely and in a slow pace. After the more demanding reading task, they were presented a slideshow containing the other target words. The whole procedure took approximately 20 minutes for each informant. Only at the end of both tasks were speakers fully informed about the purpose of this research. Approximately, only one speaker out of four was able to spot that the r-sound was the actual goal of the recording session.

8. Procedures of the analysis

8.1. Preliminaries

For each speaker five sound files are realized (one for each of the four sections of the word list and one for the cueing screen elicitation task record). Each of them contains approximately forty target words. Prior to the analysis of the sounds, two preliminary stages take place. First, the distracters present in the four speech files are cut out. Second, all the speech files are launched one by one through *Praat* and all the items of the test are segmented and labelled.

A short explanation must be provided about the second phase which is a long and painstaking procedure. Segmentation and labelling are concerned with the detection of the words and sounds we are actually interested in. *Praat* allows us to do this task through the application of borders and labels on tiers. These are blank bands located underneath the sound waves shown in the Praat sound window. On these bands, intervals are added in correspondence both to the beginning and to end of the parts of the sounds we are interested in. For this work, three tiers have been necessary.

The first one is named *type*. Its function is two-fold: it determines the portion of sound we want to analyse. In this research it defines the segment of the /r/ sound. The label on tier 1 represents the phonological context of the sound. For example, the label ˈ*yr* detects an interval whose /r/ sound falls into the context of 'stressed - long vowel /y/ in final position' e.g. *duur* 'expensive', *buur* 'neighbour'; *ɔrt* stands for 'unstressed - short vowel /ɔ/ in coda position before consonant' e.g. *kort* 'short', *storm* 'storm'.

The borders of the interval, were added at the beginning of the r-sound and at its end. During this stage of the research, we found out that in some cases the phonological slot /r/ was filled with two sounds (usually the second part was aspiration). In these cases, the position of the end border in tier 1 was set at the end of the first phonetic segment.

The second tier is named *clues*. The intervals were added freely and not in correspondence to precise points of the sound wave because, unlike the previous one, this interval is meant exclusively for adding notes or, in other words, cues on the type of sound. These cues can be provided either by the spectrogram or the soundwave or both (see sections § 8.2., § 9.1. and § 9.3.). For complex sounds (two segments for the same sound, e.g. ɻh,) the positioning of the border are not completely free (see below).

Labels on tier two have been filled only when cues were detectable by eye, otherwise it has been left blank.

Annotation on tier 2	Meaning	Consequences for the analysis
trill	The sound is judged by eye as a trill	The analysis will render either a [r] or a [ʀ], + or - voiced.
tap	The sound is judged by eye as a tap	The analysis will render a [ɾ], + or - voiced.
ə insertion	After the /r/ segment, an epenthetic ə is inserted. /r/ is not clearly definable	The analysis will render a [ɾə].
zero	The sound is dropped	The analysis will render [Ø].
2	In V/r/C# contexts only. The sound is made up by two segments	The analysis will render a compound sound, e.g. [ɹʰ]
misread	The word is misread	The analysis will not consider this word
missing	The word is missing from the sample	The analysis will not consider this word

Table 8.1. All the possible cues on tier 2.

Trills and taps are quite straightforward to spot (see next sections § 9.1. and § 9.2.). Epenthetic [ə]s are also quite easy to spot (and hear). For the sake of clarity and brevity, all the cases of [ə] epenthesis are labelled as [ɾə] because the most common sound enabling the schwa is a tap[3].

If the /r/ sound was totally dropped, the tier was accordingly labelled as *zero*. This was not as easy as it may sound. This category is represented by the cases in which the spectrograms do not provide a clear changing pattern between the preceding vowel and the end of the word. At the same time, they also give not impressionistic cues of any sounds for /r/.

If the target part of the word was misread the tier was accordingly labelled *misread*. In some cases the word was missing in the speech file (usually this happened when the speaker mistakenly skipped it on the reading list) and accordingly it was accordingly labelled *missing*.

3 Yet, throughout the sample I have come up with different sounds such as trills or non-rolled sounds. Although the script was originally designed for recognising these differences and to deliver a precise outcome, I deliberately chose to convert all these sounds into [ɾə] because I wanted to regard them all as deviating the phonological context from postvocalic to intervocalic.

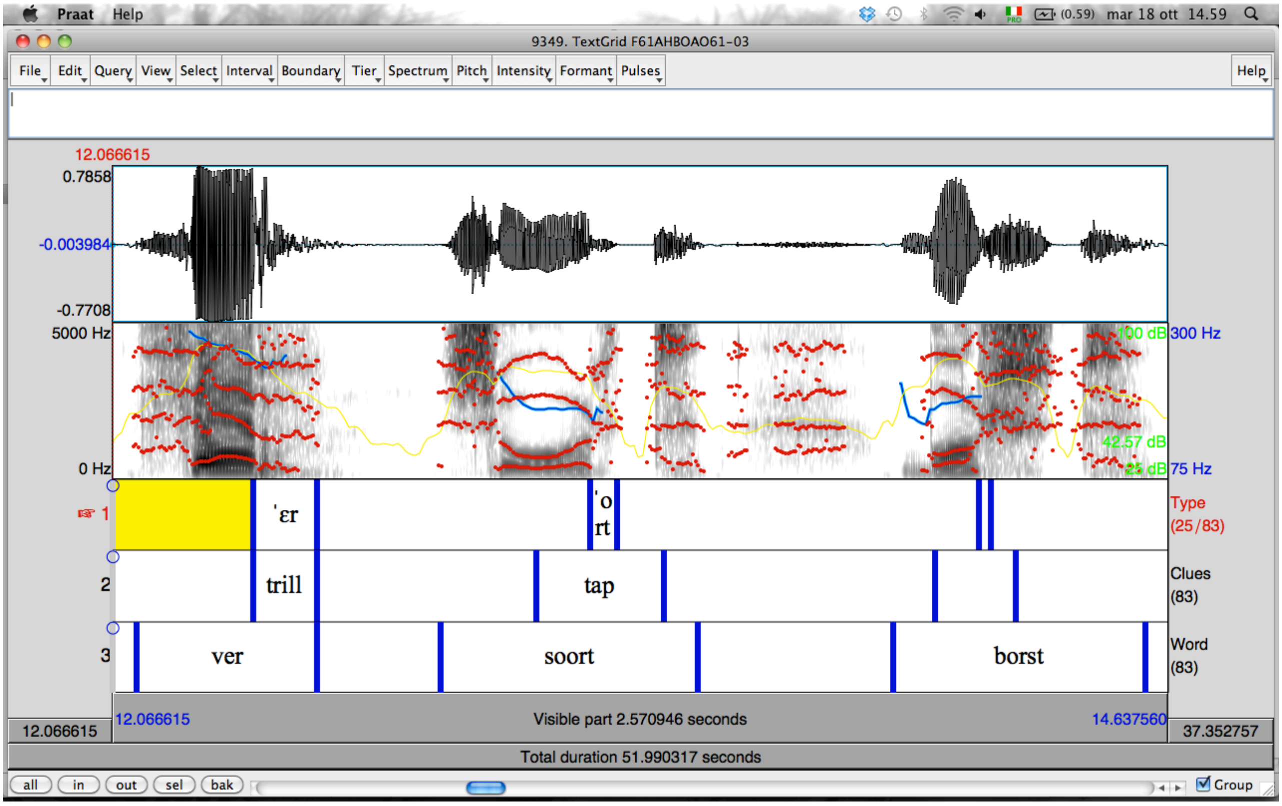

Fig. 8.2. A picture of the *Praat* textgrid window during the segmentation stage: adding borders and labelling intervals. You can see the three windows: from top to bottom: the sound wave, the spectrogram and the three tiers. On the right side you can see the names of the three tiers. No cues were found for the r-sound in *borst* 'chest', so the interval has been left unwritten for the following automatic analysis.

The last important notion about the second tier is the label '2'. It is just a conventional annotation that basically means 'this /r/ sound is made up by two main segments'. This, however, occurs only in the phonological contexts that end in C i.e. V/r/C# (in V/r/# contexts the two /r/ segments are recognised automatically by the script through the difference between the end-border of the word and the end-border of the /r/. In fact, when the /r/ sound in this context is made up of just one segment, the end-border of the /r/ segment must match with the end-border of the word (as in fig. 8.2. *ver* 'far'). If not, the discrepancy is due to either C# contexts (as in fig. 8.2. *soort* 'type') or to a two-segment sound). To summarise, if two segments for the same sound are visible in the spectrogram, in context V/r/C#, the end border in tier 2 matches with the end of the second segment. (These are the only cases, in this tier, in which a border cannot be positioned freely).

In many cases tier 2 is left blank (see fig. 8.2. *borst* 'chest'). This mean that the analysis of the sound, and its subsequently classification, could not be safely done by eye. The sound therefore will be analysed automatically through a specific script formulated for this work (section § 9.4.).

The third and last tier detects the target words and it has been named as such i.e. *word*. Therefore, the boundaries of the intervals are positioned in correspondence to both the beginning and the end of the word. Each interval in this tier has been labelled with the word recorded in the sample.

8.2. Analysis, writing (or not) clues on the second tier

Once the three tiers are filled through the application of borders and labels for the speech files of the sample, the pivotal phase of the research begins: the classification of each occurrence of /r/sounds produced by informants during the test. This is the most crucial and demanding constituent part of the whole work. In what follows, a detailed description of the steps undertaken will be outlined.

We saw in the previous section, that part of the analysis could be done by eye or automatically through a computer-based analyser. What drives the two analyses is described in this section that basically deals with the acoustic cues we need to find in order to classify the sounds from our sample.

First of all, let's see what can be useful in order to discern the sounds from spectrograms and soundwaves. The acoustic dimensions most relevant in the description and classification of consonants are, according to Hayward (2000), 'presence or absence of aperiodic energy, timing and spectral structure' (Hayward 2000:176). Taking these hints as a starting point, we need to verify whether they can be useful for the needs of this research.

8.2.1. Aperiodic and period energy in the soundwave

A first cue for discerning sounds is given by the pattern of the sound waves. This point is optimal for discriminating between, for instance, fricatives and stops and affricates. Our main interest, however, is (mostly) on rhotics sounds which most of the time are sonorant consonants, which show a periodic energy throughout the production of the sound. However, sometimes /r/ in Dutch is realized as a fricative e.g. [ʁ] [χ] (see table 4.1.). Fricatives can indeed be discriminated for the aperiodic energy superimposed in the sound wave. The harmonicity-to-noise-ratio (discussed in detail in section § 9.4.3.) will be computed for singling out these sounds.

Apart from the cases of fricatives, all other rhotic sounds show periodic soundwaves. Because of that the first point is, in general, of little assistance. To give an example, see the fig. 8.3.. It shows the soundwave representing the transition from the end of the vowel and the beginning of the /r/, taken from a word in the sample of speaker F22JBAA22. The change is very subtle and we cannot judge when the former leaves for the latter. We can only notice an increase of the sound waves in the second part, but the aperiodic energy seems utterly absent.

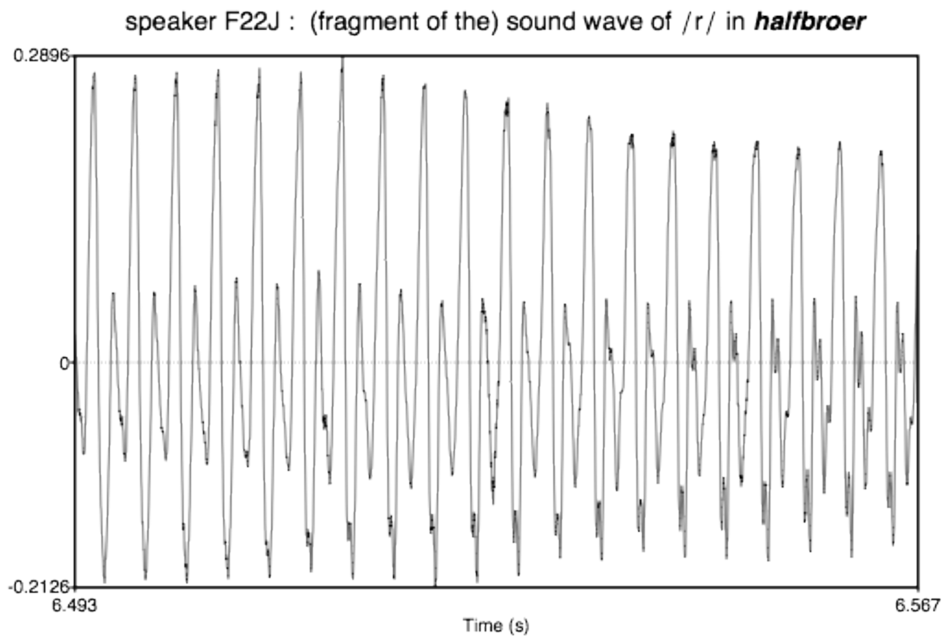

Fig. 8.3. Portion of the sound wave of *halfbroer* 'half brother' (speaker F22JBAA22) performed through *Praat* v. 5.3. Transition between [u] and /r/ show a steady periodicity.

8.2.2. Length of the sound

If the first hint was of little assistance, then length of the sound is, unfortunately, of none. Timing is completely irrelevant for the acoustic classification of sounds, of our sample. In fact, since the different natures of the sounds labelled as rhotic and the variety of sounds we expect in the /r/ sounds of our sample, this property does not help in discriminating between sounds. In the sounds gathered in the sample, duration seem totally uncorrelated with the types of sound. One simple example is provided by some /r/ durations taken from the same speakers of the sample: speaker M18AVMBOAA18 produced, among others, alveolar approximants [ɹ] (*kantoor*, 'office' 117 ms), vowels [a] (*bouwboer* 'arable farmer' 115 ms), [ø] (*acteur* 'actor' 111 ms), fricatives [x] (*hier* , 'here' 117 ms). All these values are very similar, therefore rejecting timing as a reliable source for discriminating among the consonants we were looking at.

Furthermore, a second point on the absolute unreliability of the length of the sound as an acoustic cue for the classification of (our) sounds comes from the setting of the borders: the beginning of the /r/ is, especially for approximants and vowels, unclear and blurred. These sounds, in fact, clearly compound a continuum with the previous vowel. It is worthwhile to remind that '[a] consonantal constriction does not just happen; the articulators must move into their target position and then move away to the target position for the following sound. When a vowel precedes or follows these movements will be reflected in movements of the vowel's formant from or to their steady-state values' (Hayward 2000:175). Now consider the following spectrograms in fig. 8.4 and 8.5.

The beginning of the /r/ sound in fig. 8.4 is rather simple to spot: it corresponds to the beginning of the first white band (at around 0.71 sec.). The spectrogram in fig. 8.5, in turn, does not show abrupt changes in the pattern although it is also remarkable that both F_2 and F_3 begin to fall at around 0.25 sec. However, this is too little to prove that at that very moment the realisation of the /r/ began. Rather, this is plausibly the moment in which the speaker leaves the articulation of the vowel for producing the following sound. 'These movements are known as the *formant transitions*' (Hayward 2000:175) [her italic]. Accordingly, we advance that the transition toward the consonantal sound begins at 0.25 sec. Yet, we do not know where to set the beginning of the /r/. Accordingly, the real beginning of the /r/ sound is not definable.

This is also justified by the articulatory gestures involved in the types of sound involved in our research, namely vowels (since we always deal with postvocalic sounds) and the approximant-like pronunciations which may follow. Vowels, by an illuminating definition, are sounds that are phonetically like 'an approximant formed with pulmonic air (as a rule egressively, i.e. through exhaling), whereby the airstream encounters no obstruction (neither stopping or friction) in the resonance chamber' [his emphasis] (Bussmann 1996:1274). Therefore, the movements of the articulators from vowel to approximant must be of little relevance. In this case, it is the rise and/or the contraction of the tongue which is moving but does not create contact nor friction with other surfaces in the mouth, that produces the approximant sound.

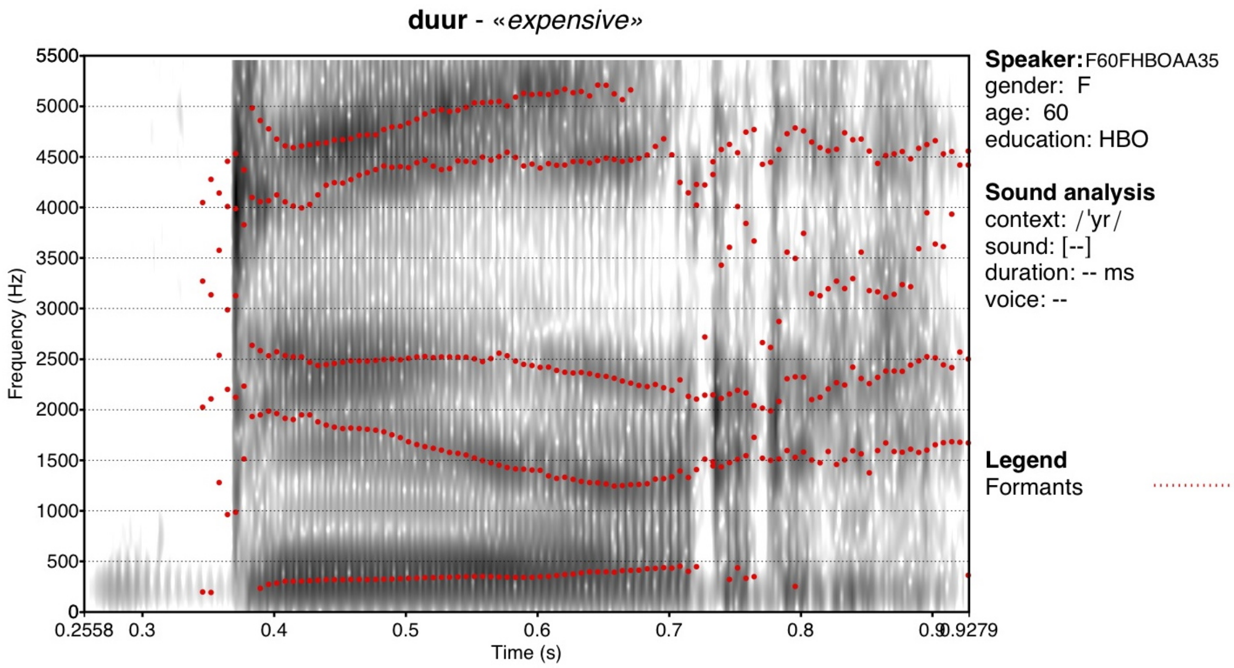

Fig. 8.4. Spectrogram of *duur* 'expensive' (speaker F60FHBOAA35) performed through *Praat* v. 5.3. The beginning of the /r/ is clear for cases like this one one, that is trills [r].

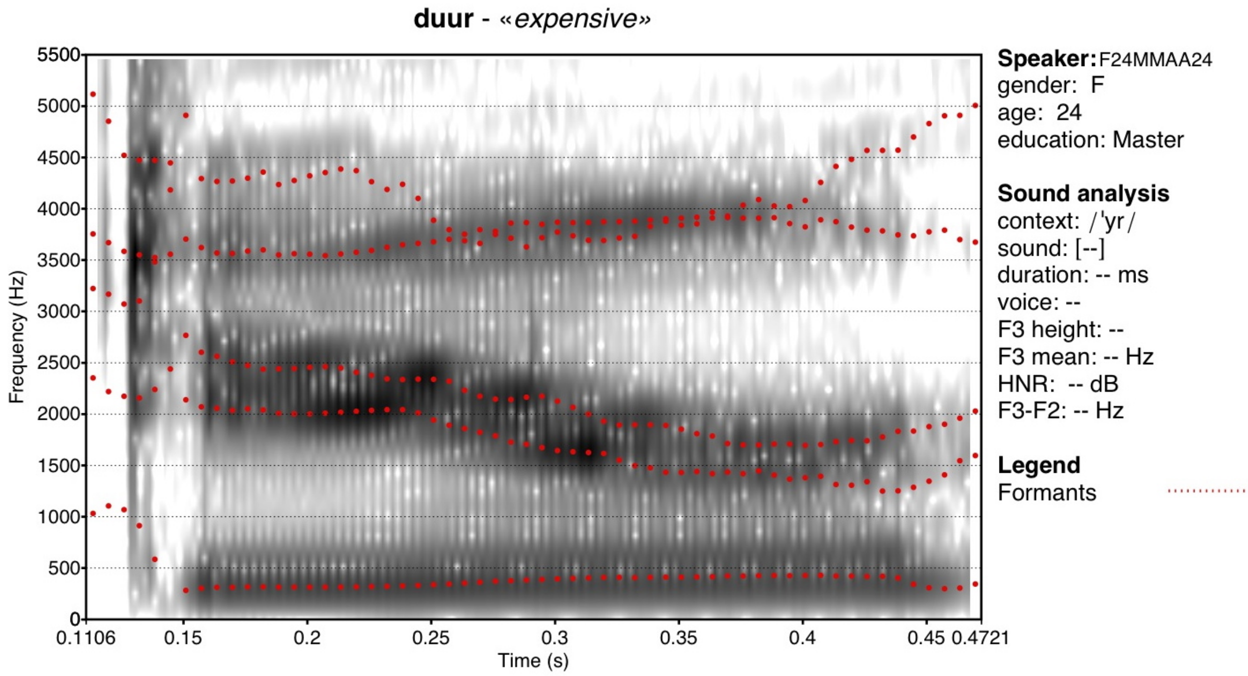

Fig. 8.5. Spectrogram of *duur* 'expensive' (speaker F24MMAA24) performed through *Praat* v. 5.3. The beginning of the /r/ is blurred with the end of the preceding vowel. The rhotic approximant shows a clear vowel-like spectral structure.

This confirms that the length of the sound is of no help because its value cannot reliably be detected in many cases. This pitfall concerns the placement of the border in the *type* tier since the beginning of the /r/ is blurred with the end of the preceding vowel. Conventionally, the border is usually set, by the eye, halfway between the beginning of the descent of the F_3 and its normalisation during the approximant /r/ sound.

Another consequence of the uncertainty of the beginning of the sound is mirrored in the computation of the formant values, which are essential for determining and classifying the sounds. Formant means are taken over the middle half of the /r/ segment. The new beginning is set at 25% and the new end at 75% of the total length of the segment.

This choice was forced by the fact that the measurements of the formant values could not be relied upon boundaries (the begin and end borders set in tier one) which are unreliable by 'definition' (because they can be blurred and mistakenly placed, especially in cases of approximant sounds). In so doing, part of the sound is discarded but at the same time the reduced sound-space would narrow the chances of wrong and deviant measurements. The solution adopted appears to be a rather convenient compensation of the two factors.

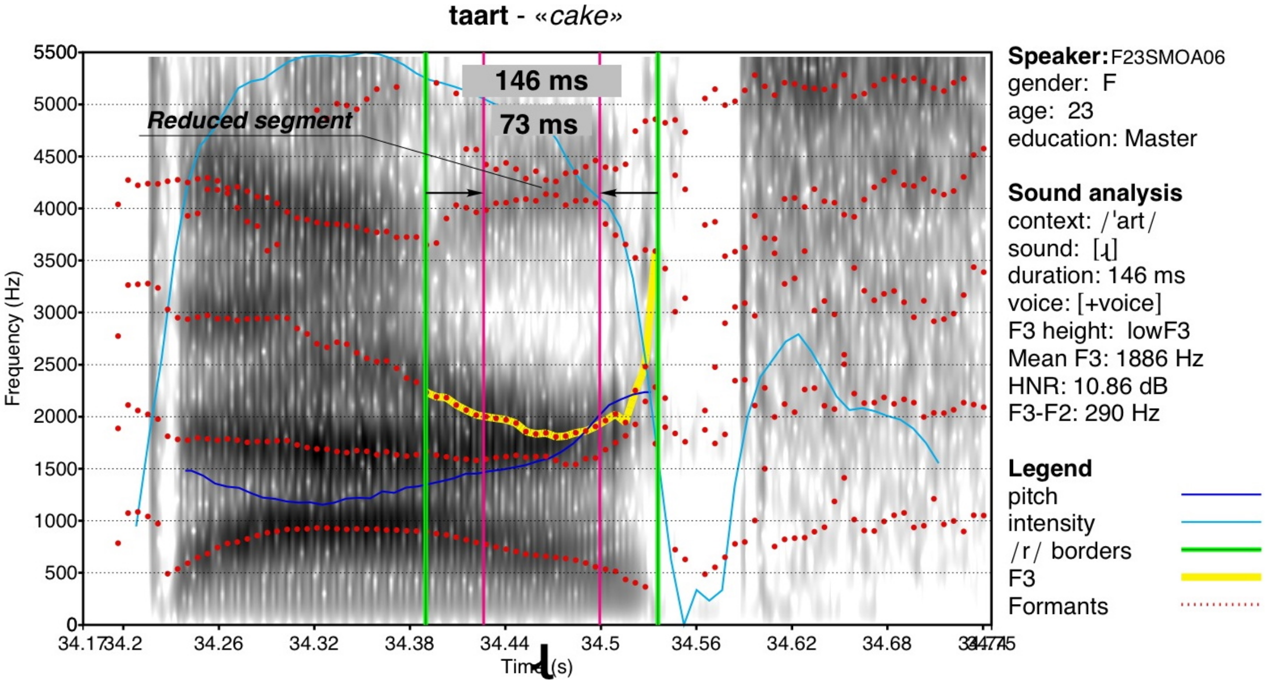

Fig. 8.6. Spectrogram of *taart* 'cake' (speaker F23MOA06) performed through *Praat* v. 5.3. Tier 2 was left empty because the sound was not a trill nor a tap (see section § 9.). The Praat script is designed to compute the phonetic measurements for a reduced segment of 50%.

8.2.3. Spectral structure

The last acoustic information, the structure of the spectrogram, is of major importance for the analysis of sounds. Through the 'reading' of the spectrograms much information can be immediately drawn. In figure 8.4, we easily spot a trill-like pattern. In many cases however, spectrograms cannot give us precise visual cues for determining and classifying the sound. In figure 8.5, the spectrogram shows a continuum and we cannot determine by eye what sound is following the syllabic nucleus. As a result, to classify these sounds, deeper insights and cues must be sought. In the following section the steps undertaken for the classification of the sounds will be thoroughly described.

9. Analysis both by eye and by computer

Among the three acoustic pieces of information mentioned by Hayward one is of no help (timing), one is of little help (aperiodic energy on the sound wave discriminates fricatives though), while the structure of the spectrograms is important for the detection of some sounds. Yet, since the inventory of the sounds produced by speakers during the test will be rather wide, some sounds cannot be found only through the observation of the sound wave and the spectrogram. For example, since a vowel is a *kind* of approximant, how can we discriminate between the two? How much aperiodic energy does a sound need to be considered a fricative? Indeed, these questions cannot be answered by eye.

More refined methods and processes are necessary to conclude a reliable classification of sounds through their acoustic properties. In what follows, I provide the description of both the analysis by eye and by computer, which together constitute the central part of the analysis.

9.1. First step: catching trills

We already noticed that some sounds are easily detected through the visual cues provided by the spectrogram. In fig. 8.4 we showed that the beginning of the /r/ segment was clearly visible as being the point in which the vowel-like spectral pattern stops and a some 30 ms white band appears. From this point onwards the pattern gives us even more indications than simply the starting point of the consonant sound. That sound in fact, is a good example of a trilled /r/, presumably alveolar [r].

Phonetically, this type of sound is produced through 'the vibration of one speech organ against another, driven by the aerodynamic conditions. One of the soft moveable parts of the vocal tract is placed close enough to another surface, so that when a current of air of the right strength passes through the aperture created by this configuration, a repeating pattern of closing and opening of the flow channel occurs' (Ladefoged and Maddieson 1996:217). On a spectrogram, this sound reveals a series of white and black bands: the former are the consequence of the closed phases and the latter of the opened ones.

The examples below (fig. 9.1 and 9.2) show two neat cases of trill sounds. The patterns precisely follow the indications available from the literature. As seen in section § 4., contemporary Dutch is featured by two types of trill sounds: the alveolar realisation [r] and the uvular one [ʀ]. How to discriminate between the two points of articulation is also rather straightforward from the observation of the spectrogram's aspect: for [r] the closed phase is longer compared to [ʀ] and the total number of vibrations is lower in the former. However, the examples above are clearly prototypical. On some other occasions neither the spectrogram pattern nor the impressionistic evaluation give safe assessments about the point of articulation of the trill. Because of that, in tier two the cue is just *trill* and the point of articulation will be automatically drawn.

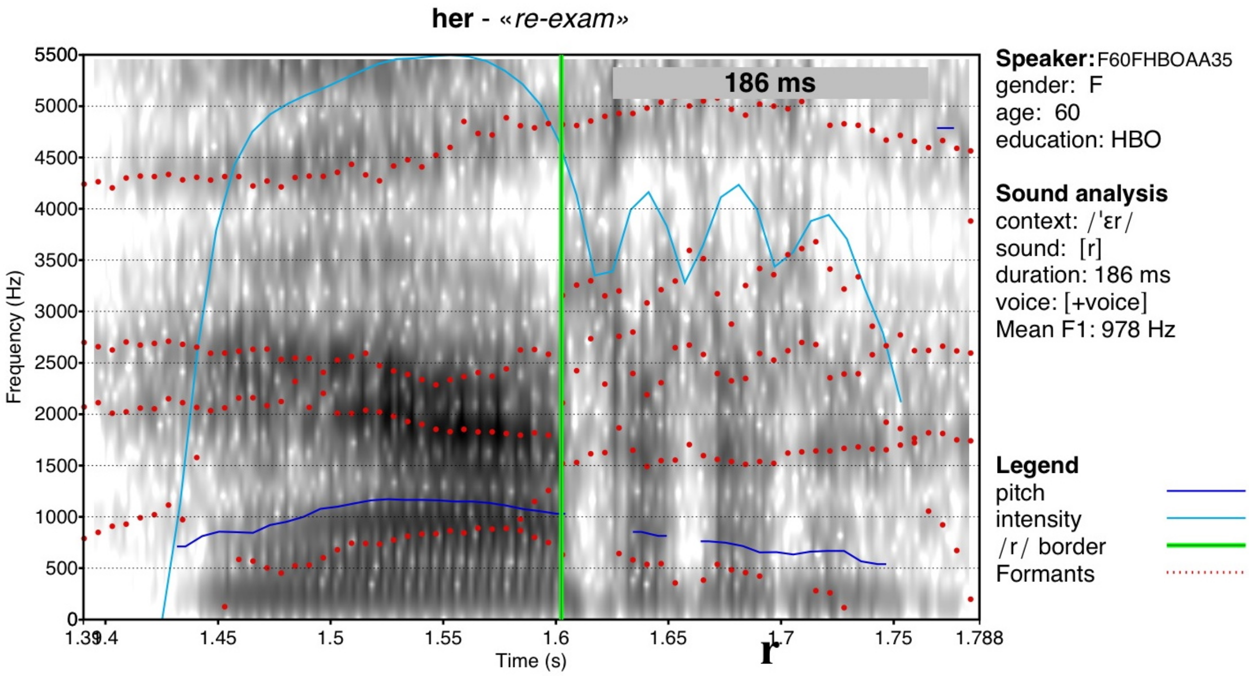

Fig. 9.1. Spectrograms of *her* 're-exam' (speaker F60FHBOAA35) performed through *Praat* v. 5.3. Instance of alveolar trill [r].

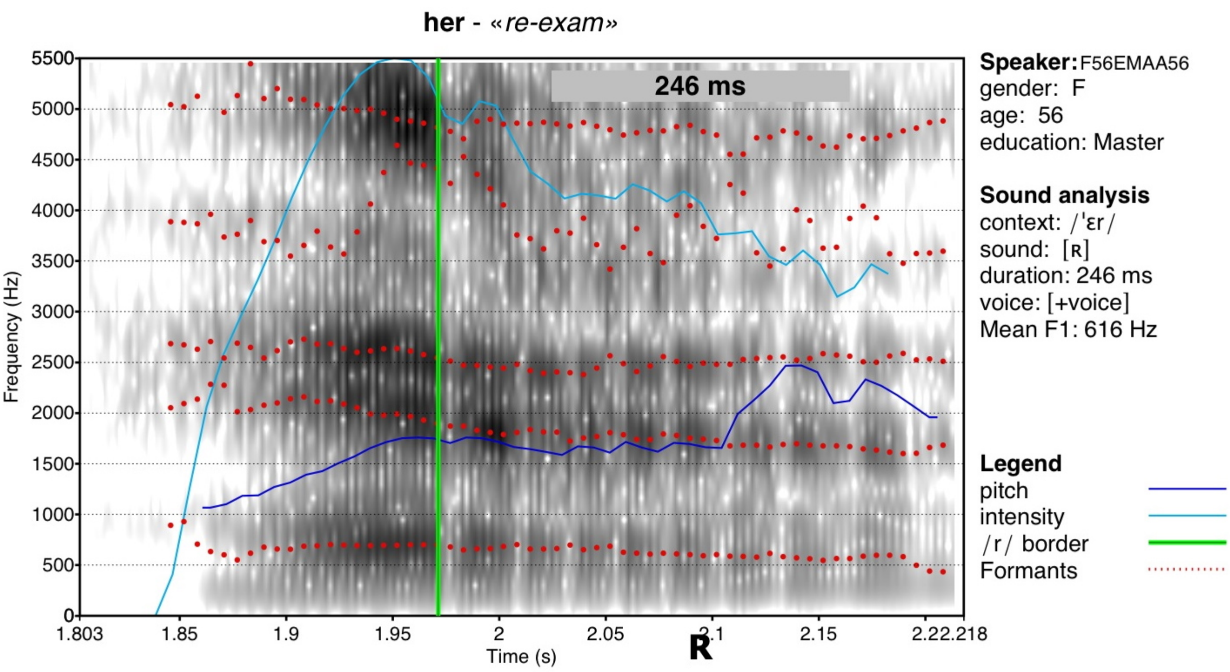

Fig. 9.2. Spectrograms of *her* 're-exam' (speaker F56EMAA56) performed through *Praat* v. 5.3. Instance of a uvular trill [ʀ].

Although it has been pointed that the third formant is a reliable factor for discriminating between the two different articulations, 'there is a consistent distinction in the spectral domain between uvular and apical [alveolar] trills, with the uvular trills showing a much higher third resonance […]' (Ladefoged and Maddieson 1996:227) the present analysis does not takes into account this path. Apparently, the third formant resonance is highly unreliable for the data of this work. For example, the F_3 means in the two spectrograms above are in fact very close together with even a higher values for [r] 2,670 Hz against for [ʀ] 2,439 Hz. An alternative way of discriminating between the two sounds has been developed through the modelling of the analysis *Praat* script of this research. The attention has been focussed on the F_1. While in the case of [ʀ] the red dots highlighting the formant seem to have a stable path, in the case of [r] they completely disappear in correspondence of the closed phases. This is a side-effect of the closed phase duration difference. *Praat* extracts the values of the formants from windows of sounds measuring 25 ms each (see § 9.5.1.). Since, typically, the closed phase in uvular trills is shorter than this value, *Praat* does not catch the *real* F_1 in that small segment. On the other hand, since a typical closed phase produced by an alveolar trill is longer than the analysis window, Praat renders a reliable, much higher than the previous one though, F_1 measurement.

This discrepancy seems to cause the F_1 to have higher mean value for alveolar trills. A conventional cut-point has been detected in 900 Hz: what is above this threshold is considered an alveolar trill, what is below, a uvular trill. In the examples above we discover that 1,001 Hz is the average value for the F_1 of the alveolar sound whereas 612 Hz is the value for the uvular one).

A comment must be added here: timing, at least in this particular context, does help us in discriminating between sounds. However, what has been said in § 8.2.2. and § 9. still holds since duration cannot give us hints on the manner of articulation but, instead, help us in discriminating points of articulation within a small set of sounds, namely trills.

9.2. Still the first step: catching taps

Although similar to the class of trills, tap sounds are classified on their own because the articulatory organs produce no more than one vibration. They are also named *one-trill sound* to call attention to their similar nature with trills. The short closure leaves the expected white band in the spectrogram and no more vibrations can be observed but, instead, a noisy fricative-like sound possible due to a friction during the phase in which the tongue is retracted to reset to a steady state.

Two different points of articulation, according to the IPA consonantal chart, can be involved in the production of a tap sound. If the tip of the tongue hits the area just above the teeth an alveolar tap is produced [ɾ], otherwise, if it hits the roof of the mouth through a more retracted movement a retroflex tap is the result [ɽ]. Unlike for trills, here the difference shown by spectrogram patterns seems less obvious. Accordingly, both taps will be expressed with [ɾ] regardless of the point of articulation. The cue on tier two will be just *tap*.

It must be stressed that although taps (and trills too) are described in the literature as voiced sounds, in many cases the sample provides examples of unvoiced sounds (like those in fig. 8.7 and fig. 8.7). The unvoiced trait is displayed by the symbol ̥ superimposed underneath the IPA symbol.

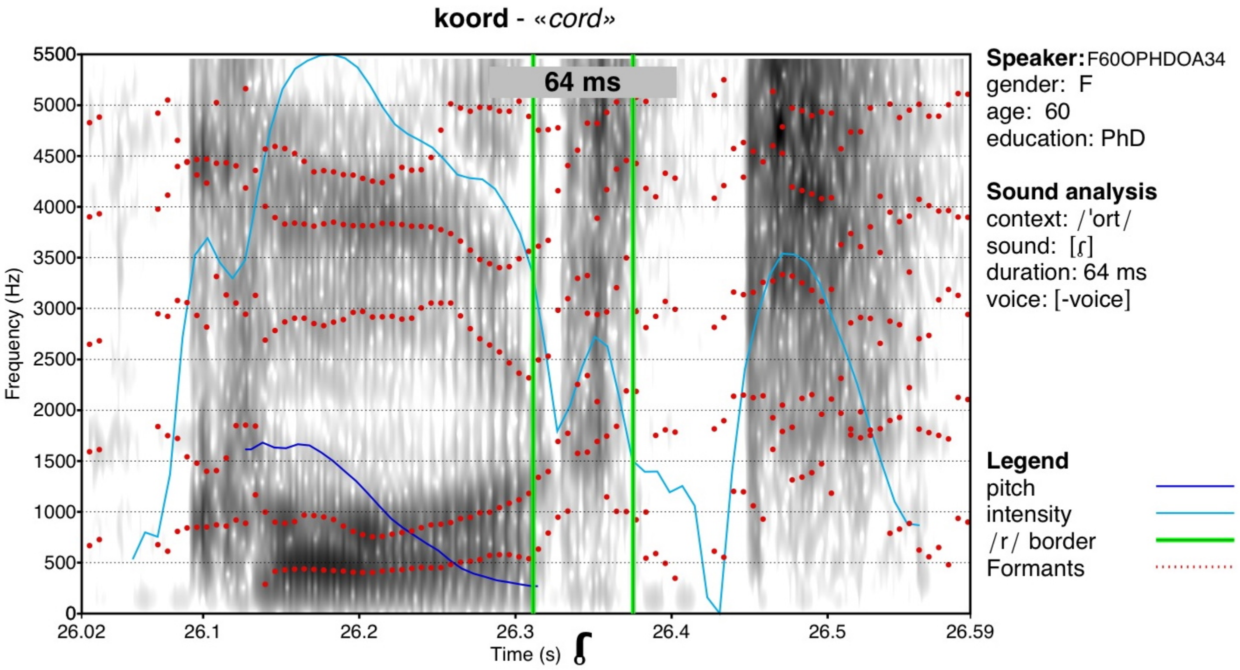

Fig. 9.3. Spectrogram of *koord* 'cord' (speaker F60OPHDOA34) performed through *Praat* v. 5.3.

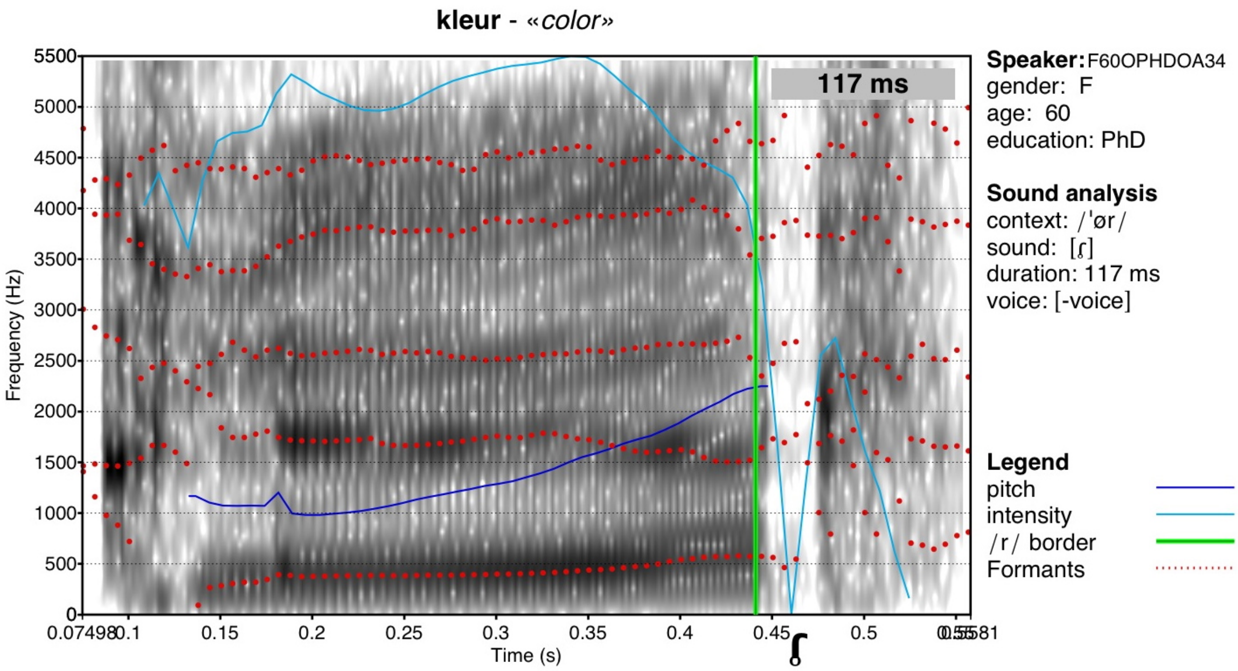

Fig. 9.4. Spectrogram of *kleur* 'colour' (speaker F60OPHDOA34) performed through *Praat* v. 5.3.

The length of the closed phase may considerably vary between speakers and within the same speaker. (The phonological structure may play a role, but this is not studied here). The second example in fig. 8.8, *kleur* 'color' shows a tap whose closed phase lasted 30 ms, more than double of the one in *koord* 'cord', in the previous example. The total lengths, 64 ms and 117 ms, of the /r/s gather the closed phase and the friction-like noise produced by the retraction of the tongue. Especially in *koord*, you can see that the green borders (representations of the interval set in tier 1) are set as follows: the beginning intersects the end of the vowel with the abrupt start of the tap and the end matches with the beginning of the silent phase namely the beginning of the following stop consonant. This proves that the noise spanning from the end of the closed phase still 'belongs' to the tap.

9.3. Still the first step: other cases

9.3.1. Epenthetic [ə]

A very common feature of the contemporary Dutch is the insertion of a schwa [ə] in phonological contexts such as /rk#/, /rp#/, /rm#/, /rn#/, /rf#/ and /rx#/ e.g. *werk* 'work' /ʋɛrk/ > /ʋɛrək/. We can discern this phenomenon quite easily through the pattern of the spectrogram. Firstly, because usually the /r/ is produced as a tap, so a white band is visible in the pattern. Secondly, because the insertion itself is quite easy to hear and see in the spectral structure. Fig. 9.5 is a good example of the epenthesis. These cases will be catalogued under the label *ə insertion* in tier 2 and the outcome of the analysis will always be [ɾə] (see section end of section § 6.1.); their occurrences will be counted up and tested to see whether a correlation between this phenomenon and the approximant pronunciation of the postvocalic /r/ is emerging (§ 10.2.6).

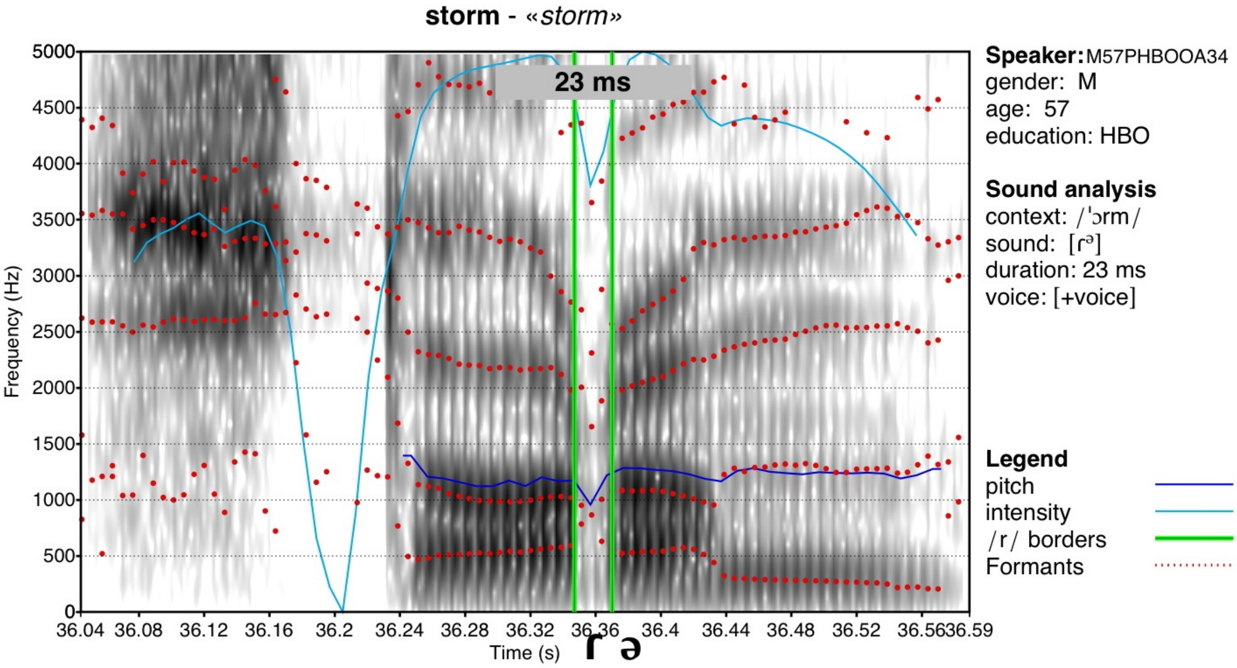

Fig. 9.5. Spectrogram of *storm* 'storm' (speaker M57PHBOOA34) performed through *Praat* v. 5.3. The single white band measures 23 ms and represents the closed phase occurred when the tip of the tongue against the anterior palate produces therefore the tap sound. Before the articulation of /m/ there is the epenthesis of a short vowel (around 60 ms) resembling a [ə].

This phenomenon is too crucial to be overlooked because it changes the phonological context from postvocalic to intervocalic (V/r/C > V/r/VC). Crucially, the phonological environment is modified and interestingly, '[the epenthetic schwa] precludes the use of the front approximant variant' (van Bezooijen, Kroezen et al. 2002:4).

9.3.2. Zero sound and special cases

If the sound that should cover the phonological /r/ slot is not present, tier 2 will simply indicate *zero*. The symbol adopted in the picture is [Ø]. The expected consequence for the dropping of the sound is the lengthening of the preceding vowel. The vowel /ʏ/ of the word in figure 9.6 measures approximately 0.165 sec. that is more than the usual length of the vowel (the length of the same preceding vowel from a survey of six randomly selected speakers of the sample spans from 0.065 to 0.125 sec.).

If a word is misread (that is, the phonological context is not respected because of a reading mistake) the sound will not be analysed. Accordingly, tier 2 will be filled with *misread*. The last case, the word is missing from the sample (mistakenly skipped during the reading test by the speakers or missing from the reading material because of slight changes in between the recording sessions). In tier 2, thus the cue is *missing*. (Tier 1 had to delimitate a fictitious /r/ segment and tier 3 a fictitious word[4]. Straightforwardly, no symbols are adopted for the last two cases).

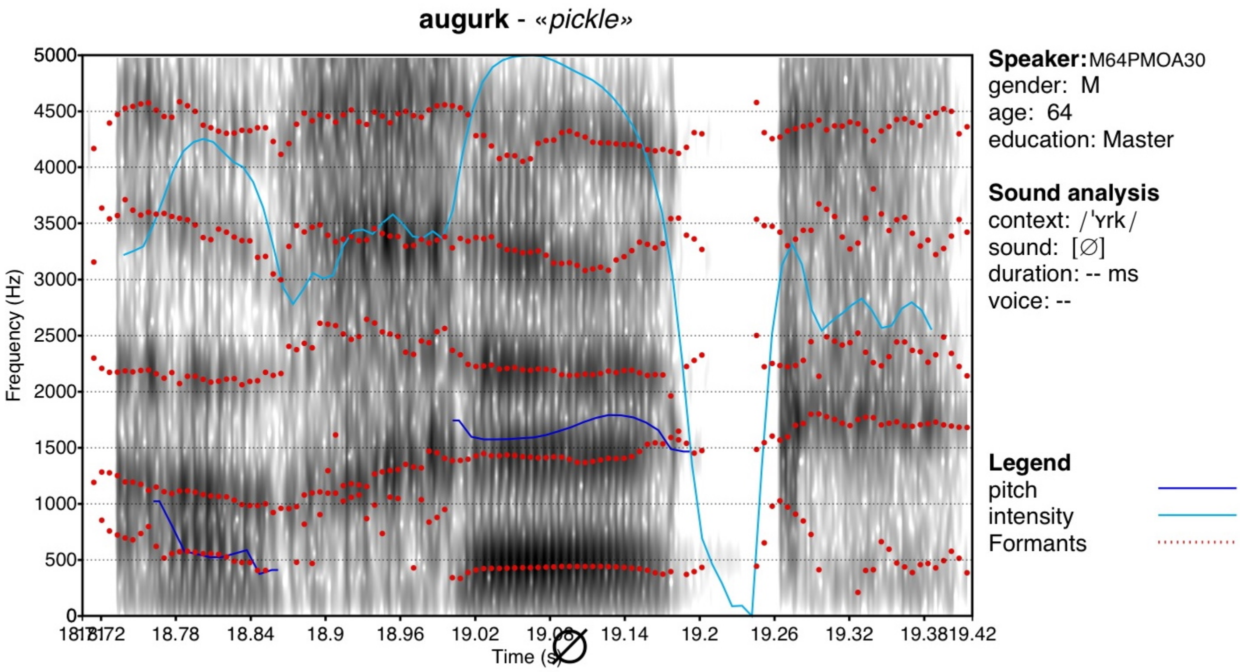

Fig. 9.6. Spectrogram of *augurk* 'pickle' (speaker M64PMOA30) performed through *Praat* v. 5.3. There is no trace of any sound filling the expected phonological /r/ space. The zero symbol is placed just below the vowel which shows a steady pattern until the beginning of the silence signalling the upcoming plosive consonant.

9.4. Second step: no visible cues; the three problems faced by the analysis

Obviously, if a sound is not a trill nor a tap it must be something else. Since secure cues cannot be drawn through the spectrogram itself for sounds beside rolled /r/s, tier 2 is left blank because an automatic analysis run through *Praat* 5.3 will draw the information needed to classify the sound. This was, probably, the most high-demanding task of the entire postvocalic /r/ project, that is the elaboration of a computer program that analyses and classifies the sounds. Three major problems arose from three different angles for this undertaking:

A **conceptual** problem. Even though we can expect, at least, the ten sounds listed by Van de Velde (1996) (table 4.1), we can also expect, theoretically, any possible sound for the /r/ segment. In other words, an ideal program, still theoretically, should be able to pinpoint any sound from those listed in the IPA chart: that is, any sound considered possible for the human body to produce. This, however, is far from being easy and, for this work, would have meant an excessive amount of work.

Because of that, the program will classify a limited set of sounds: trills and taps are given (but it discriminates between alveolar and uvular trills and between voiced and unvoiced sounds),

4 This is not a useless exercise. In so doing, for each speaker the classification table will be identical, displaying 210 items, regardless of their actual analysis.

approximants (section § 9.4.1.), vowels (section § 9.4.2.), and fricatives (section § 9.4.3.) are classified by the script. Nonetheless, these are the more common realisations of the sound taken in analysis. We consider, therefore, that the sounds most likely to arise from the sample are those listed by the van de Velde (1996) study[5].

A **phonetic** problem. This is possibly the hugest obstacle: how can we discriminate between sounds? What are the acoustic cues we need to look at to draw the classification? What are the cut-off points that make a sound fall into a category instead of another one?

A **practical** problem. How can we convert the answers to the previous questions into a computer language that is, *Praat* scripting? The following sections deal with all these problems, explain the choices and describe the settings for the analysis of the sounds. At the very end of the entire section, a flowchart (fig. 9.13.) summarises the steps of the phonetic analysis and classification of sounds of this research.

9.4.1. Approximants

Approximant sounds are possibly the main target of the entire work. We want to discover in fact whether the postvocalic /r/ is nowadays more frequently pronounced as such in the D.o.A..

The term approximant refers to the manner of articulation of the sound 'in that one articulator approaches another, but the degree of narrowing involved does not produce audible friction' (Crystal 2008:32). An approximant may be described as the opposite of a plosive sound. In fact, in the latter, there is a full contact between the speech organs whereas in the approximant-manner domain there is not any closure of the vocal tract.

5 Slightly differences are based on different classification of the sounds.

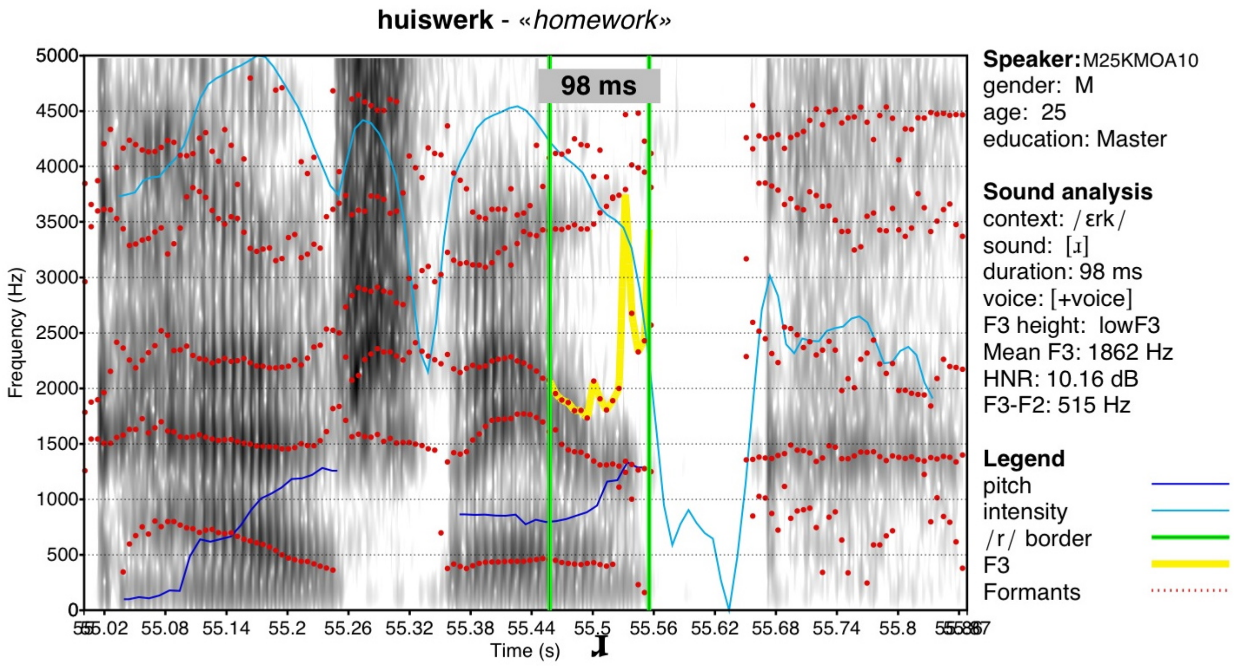

Fig. 9.7. Spectrograms of *huiswerk* 'homework' (speaker M25KMOA10) performed through *Praat* v. 5.3. F_3 is highlighted in yellow for the duration of the /r/-segment. The mean value (taken from ¼ to ¾ of the duration of the sound) is 1,862 Hz. The F_3-F_2 difference exceeds 450 Hz.

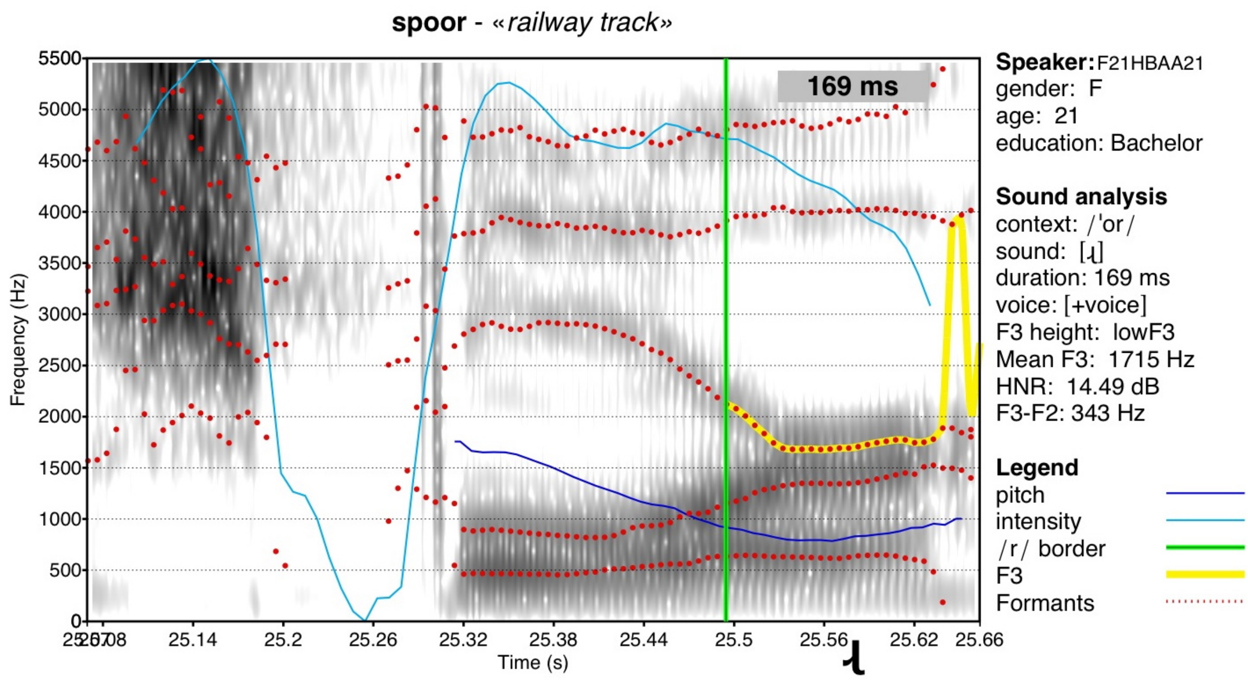

Fig. 9.8. Spectrogram of *spoor* 'railway track' (speaker F21HBAA21) performed through *Praat* v. 5.3. The F_3 moves downwards whereas the F_2 moves upwards. During the production of the sound the two formants get very close between each other. When the difference of the two means (computed considering the space beginning at ¼ and ending at ¾ of the total duration) is below 450 Hz the approximant is considered retroflex [ɻ].

This feature makes them quite related to vowels: '[c]onsonants generally involve the contact or near-contact of large areas of the articulators, whereas for vowels the articulators do not make any type of close contact. Consequently, the manner of articulation is *always that of an approximant*' (Collins and Mees 2003:61) [my emphasis].

Among this class of sounds, which contains [ʋ, ɹ, ɻ, j, ɰ, w, ɥ], '[t]he [English] sound **r** is more difficult to describe partly because different speakers make it in different ways. It usually involves some raising of the tip of the tongue toward a point on the roof of the mouth well behind the upper front teeth' (Ladefoged 2001:52). Nonetheless, the acoustic of these sounds seems to suggest a particular feature that can single them out from the rest. Acoustically, the approximant [ɹ] is featured by a very low third formant. We take on the hint provided by Ladefoged which states that '[w] henever there is a **r** [English. i.e. in IPA [ɹ] or [ɻ]] in a word the third formant will be below 2,000 Hz […] sometimes falling to low as 1500 Hz' (Ladefoged 2001:53).

This is crucial for the analysis of the data of this work and it constitutes the first constraint for the analysis: whichever sound shows an F_3 mean below 2,000 Hz is considered as a rhotic approximant. The discrimination between approximant alveolar [ɹ] and approximant retroflex [ɻ] is unfortunately more difficult. The latter is produced with the tip of the tongue curled back below the roof of the mouth. This movement is, however, rather difficult to define through acoustic cues.

However, since approximants are one of the major targets of this work, the discrimination between the two points of articulation has been pursued. An attempt has been made by analysing through *Praat* the spectrograms of the these two sounds downloaded from the IPA chart[6]. The two spectrograms for the two realisations in postvocalic position immediately present a discriminating factor: in the retroflex sound F_2 was stable and very close to the F_3 almost reaching it. The two formants thus are parallel and very close to each other. This configuration was not present in the spectrogram of the alveolar sound which presented, in turn, an unstable F_2.

From this finding, the idea for discriminating between the two sounds was to consider the nearness between F_3 and F_2 as the decisive factor. According to my findings, any approximant sound presenting a F_3-F_2 difference below the value of 450 Hz is considered [ɻ] (fig 9.8) otherwise an [ɹ] (fig. 9.7).

This is, however, a rough method. Given the lack of information in the literature we can even infer that the discrimination between the two sounds with only acoustic data might not be possible. According to Hamann 'only one stable acoustic characteristic for retroflexion in the formant transitions can be manifested, namely a lowered F_3, since the F_2 seems largely context-dependent' (Hamann 2003:60). I also received from Sebregts: 'I don't think there is an absolute measure in terms of the distance between F_2 and F_3 (or absolute F_3 lowness) that would allow you to

[6]Available online at http://www.paulmeier.com/ipa/consonants.html

use as a cut-off point (as in: above that point is alveolar approximant, below that point is retroflex or bunched)' (Sebregts 2011).

9.4.2. Vowels

If you recall the definition of approximant consonant, you may already guess what the constraint that makes a sound a vowel is. We saw that a sound showing an F_3 value below 2,000 Hz is a rhotic approximant. Therefore, if this value exceeds the threshold, that sound must be a vowel or a non-rhotic approximant sound e.g. [j].

The F_3 mean value is however not enough to classify a vowel precisely. First, we need to exclude fricative sounds from the options (see next section) and, eventually, plot the F_1 and F_2 to detect the quality of the vowel. Unlike consonants, vowels can be reliably singled out through their striking characteristics consisting of 'the remarkable correspondence between plots of the frequencies of the first and second formants (F_1 and F_2) and the traditional vowel quadrilateral' (Hayward 2000:147). Nonetheless, the degree of error is far from being annulled. In order to set the values of F_1 and F_2 for detecting the quality of the vowels, new speech data are drawn from the IFA corpus of the *Institute of Phonetic Sciences* of Amsterdam[7]. Among the speech files available some are focussed on the pronunciation of Dutch vowels. These files contain a series of meaningless words whose phonological structure is /hVt/. All the vowels present in the Dutch phonological systems are contained in this database. Two speakers per gender were randomly selected. The formant frequencies of each vowel was averaged. The results determined the constraints set in the scripts.

7 Available on line at http://www.fon.hum.uva.nl/IFA-SpokenLanguageCorpora/IFAcorpus/

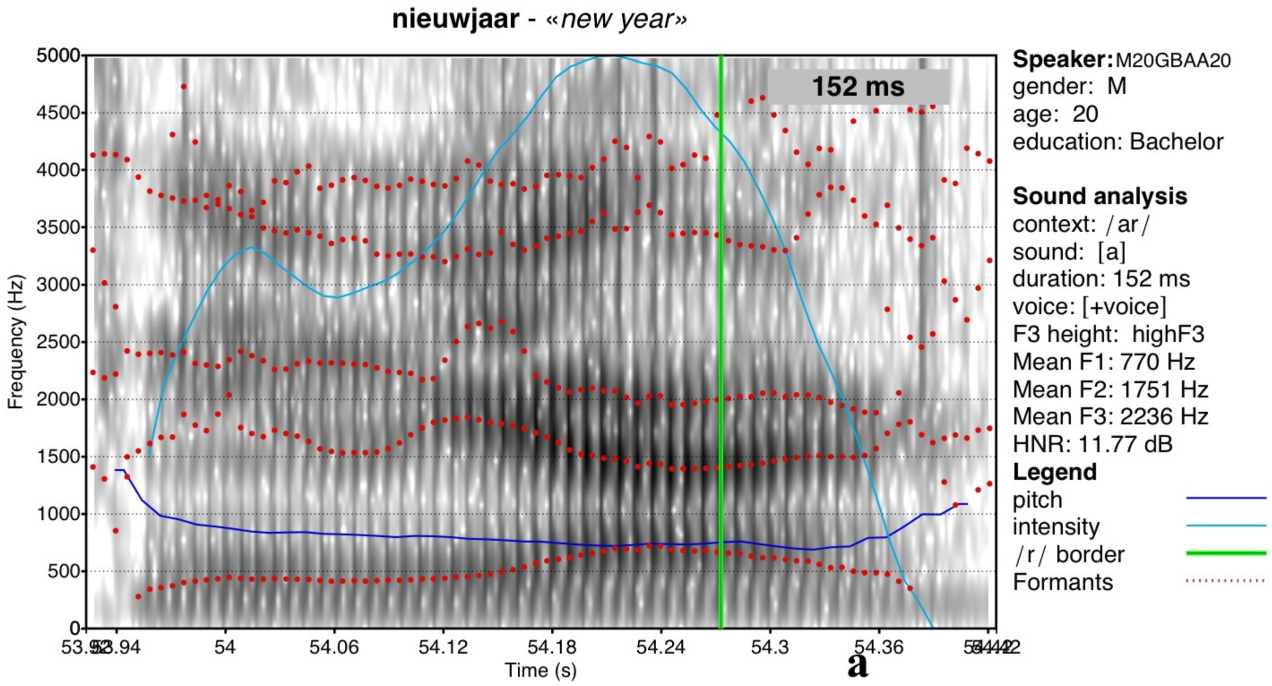

Fig. 9.9. Spectrograms of *nieuwjaar* 'new year' (speaker M20GBAA20) performed through *Praat* v. 5.3. On the right all the details that signal the sound as a vowel: a voiced sound with a high value of F_3 and a high value of the HNR. The correlations of F_1 and F_2 brought to the result of an [a]. F_3 height is labelled as 'high' that is above 2,000 Hz.

For the sake of clarity however, all the vowels classified throughout the sample were added up under the label, that is *vowel*. An exception was made for [ə] which is considered on its own.

9.4.3. Fricatives

Eventually, fricatives represent the sounds of the last category we should face throughout the classification of sounds. How can a fricative sound be separated from non-fricative sounds? The answer was already outlined in § 8.2.1 and it hinges on the aperiodic energy on the sound wave.

Trills and taps are out of question because they were detected by eye. Approximants are a *type* of vowels and *none* of them present a trace of friction. This forces us to set a cut off point based on the level of friction (aperiodic energy) to separate fricatives sounds from non-fricative sounds.

The degree of friction (noisiness) produced by the organs involved in speech production is measurable through the harmonics-to-noise ratio (HNR). '[It] quantifies the relative amount of additive noise in the voice signal. Additive noise arises from turbulent airflow generated at the glottis during phonation. […] The resulting friction noise is reflected in a higher noise level in the spectrum. Noise in the signal may also result from aperiodic vocal fold vibration. The ratio thus reflects the dominance of harmonic (periodic) over noise (aperiodic) levels in the voice, and is quantified in terms of dB'. (Ferrand 2002:481). 'The HNR is the ratio of the magnitude of the periodic signal to the magnitude of the aperiodic signal, [...]. Basically, the greater the noise component of the speech signal, the lower the ratio is' (Heffernan 2004:77)

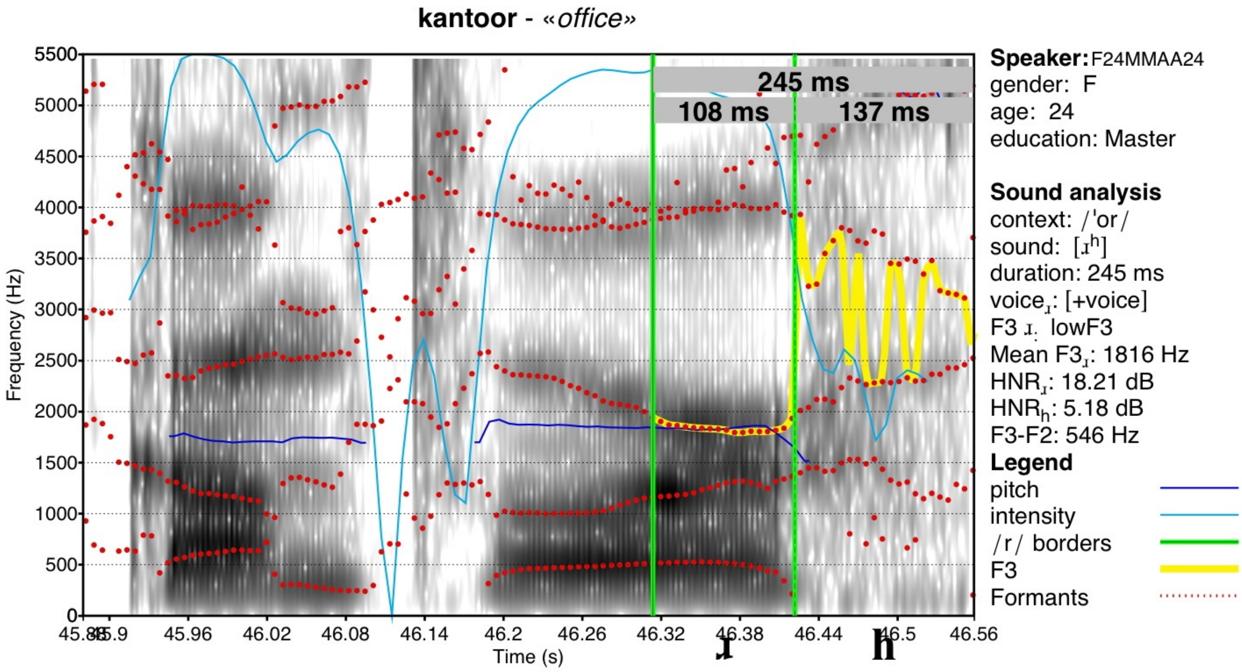

Fig. 9.10. Spectrograms of *kantoor* 'office' (speaker F24MMAA24) performed through *Praat* v. 5.3. On the right, among other details, the values of the HNRs for the two segments of the /r/ sound.

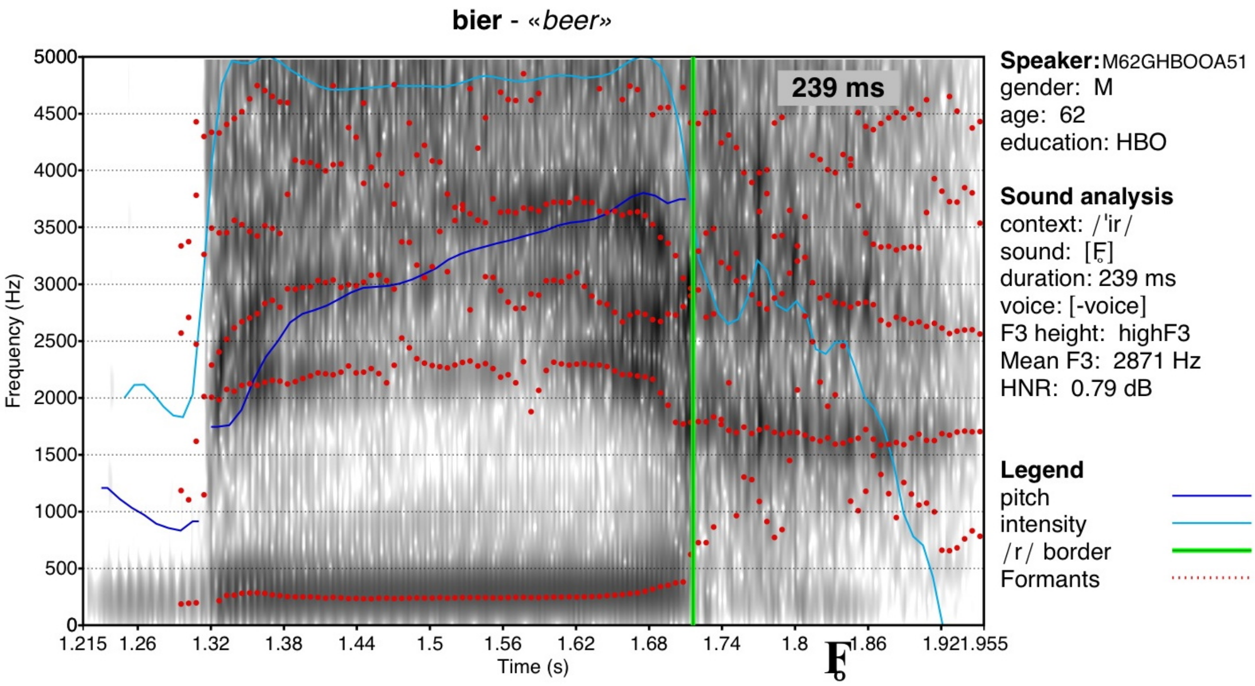

Fig. 9.11. Spectrograms of *bier* 'beer' (speaker M62GHBOOA51) performed through *Praat* v. 5.3. On the right, among other details, the values of the HNRs.

The value that sets the watershed fricative/non-fricative is fixed at 8 dB. This value has been set mainly in accordance with the data obtained through this research where both vowels and approximants show a rather high HNR between 10 and 20 dB on average. The following examples will clarify this point.

HNR values are given on the right hand of the two pictures shown above. Fig. 9.10. presents an /r/ made up by two components, the main one is the first part of the sound which has been analysed as an alveolar approximant. The second part is indeed an aspirated sound and its HNR is neatly low. The second figure represents a fricative occurrence of the /r/. The HNR is very low, less than 1 dB, and the analysis gives a unvoiced fricative.

In the fricative sounds domain a discrimination among the several points of articulation was not possible. Because of that, it must be made clear that *all* fricative sounds are signalled with a unique conventional symbol [F]. If the fricative occurs at the end of a previous rhotic element, like in fig. 9.10. in which the /r/ is made up by two segments, the symbol is [h] because, most of the times, that fricative is a unvoiced glottal.

Although several attempts were made to index the sounds based on to different phonetic cues the results were highly unreliable. This leads to a more approximated classification but it prevents from large misleading errors.

9.4.4. Creaky voice type of phonation

The normal type of voicing, also known as *modal voice*, may deviate in some occasions into a different phonation called *creaky voice*. In this kind of speech production the vocal folds are maintained rather tightly 'but not so tightly as to make vibration impossible' (Hayward 2000:262). The acoustic of the sound is predominately affected by this setting of the glottis, showing, among other incidents, very low values of HNR and unvoiced segments where we would expect voiced segments. Since it appeared evident that this occurred many times in the recorded speech[8] of the participants, we faced the problem of adjusting HNR and F_3 values across the setting of the analysis script.

The solution was to switch the prior hierarchy of the constraints in the analysis. In fact, using as first constraint the F_3 value mean was essential to prevent this error: any sound (apart from trill and taps), *regardless* of the HNR value, whose F_3 is below 2,000 Hz, is an approximant whereas anything above can be either a vowel or a fricative.

Although approximants have a lot in common with vowels, and a first separation between fricatives against non-fricatives looked apparently more obvious (this is actually what was done in the earliest version of the analysis script). However, the creaky voice phenomenon showed up in several words of the sample forcing to change the hierarchy of the constraints.

8 Apparently it occurred more often towards the end of each section of the word list.

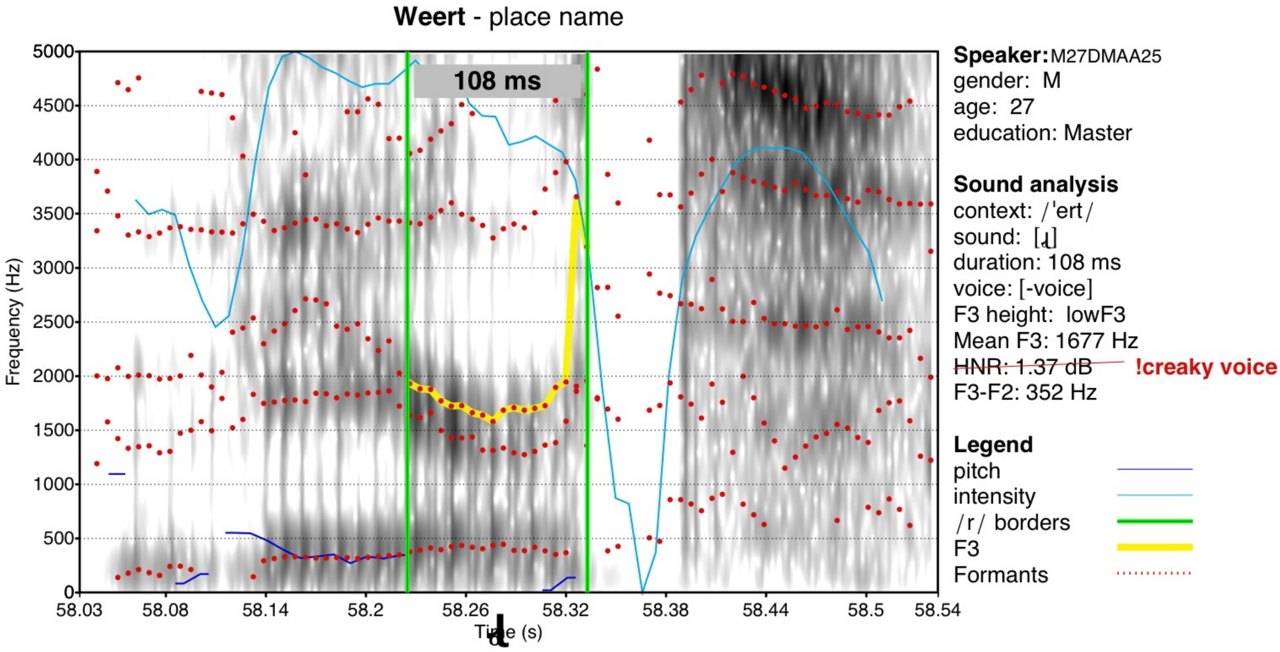

Fig. 9.12. Spectrograms of *Weert* 'place name' (speaker M27DMAA25) performed through *Praat* v. 5.3. The situation of creaky voice is 'visible' through the very low value of HNR and the lack of the pitch line for most of the duration of the /r/ segment. The sound is an approximant because of the low F_3 mean value.

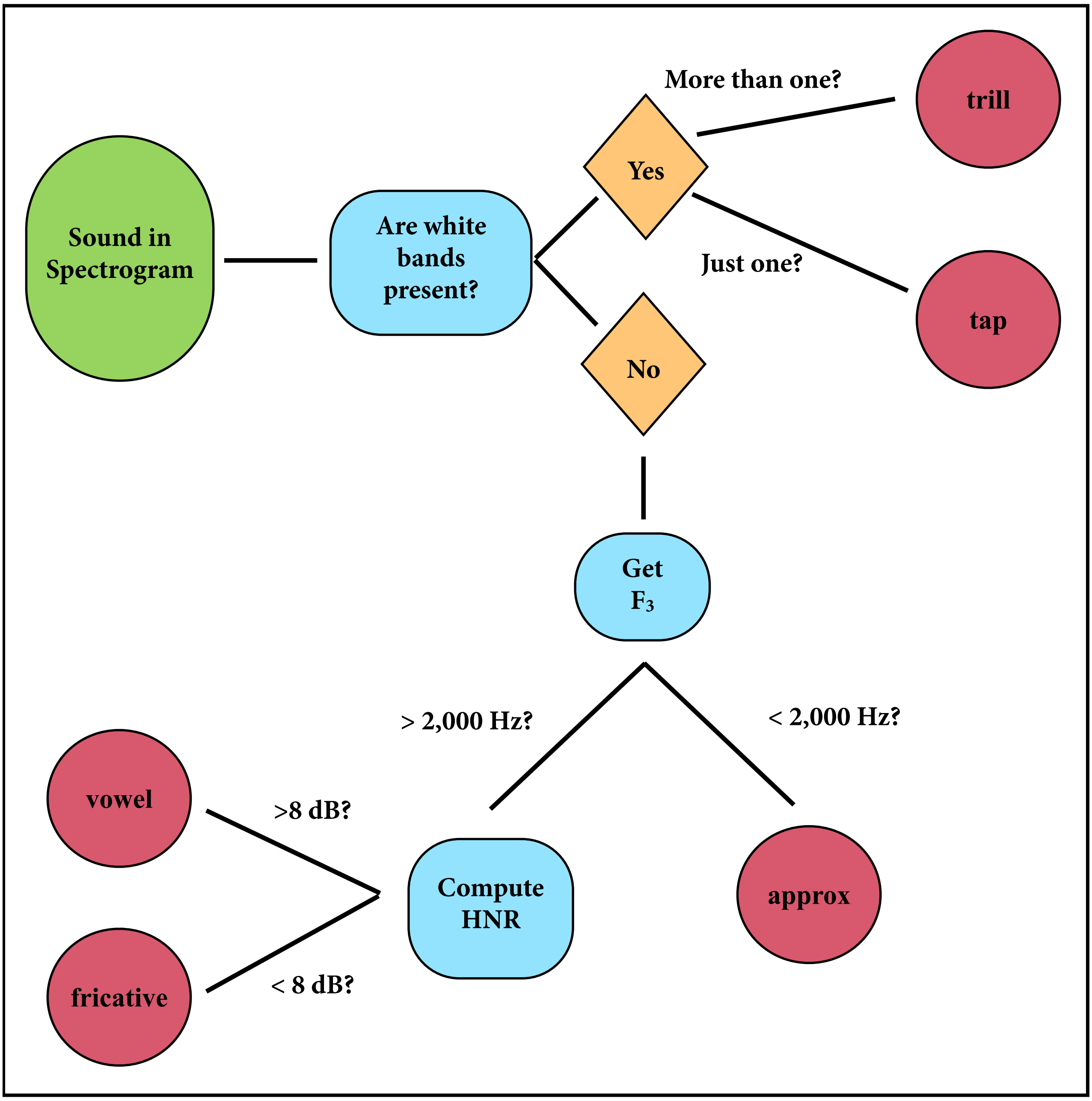

Fig. 9.13. Flowchart of the steps undertaken for the analysis and classification of each /r/ sound of the sample. The first part was carried out by eye (through the insertion of the appropriate label on tier two in the TextGrid files in *Praat*) whereas the following ones (starting from the *No* answer to the first question "Are white...") were automatically overtaken through the *Praat* script developed for this work.

9.5. Building the automatic analysis: the script and its settings

What has been described in sections § 9.3. and § 9.4 was meant to solve, at best, the **phonetic** problem. At this point it is the **practical** one that must be tackled.

The classification of the sounds is based on a computer program designed for the purpose of this research. It is a script file than runs in *Praat* and, according to the cues in tier 2 and on the specific setting dictated from the acoustic properties of the sounds, it renders for each speaker the full list of the sounds for each /r/ for each words of the speech material.

It is a complex script written in the *Praat* scripting method. I suggest to the interested reader to download the software, available for free on line at www.Praat.org, and read the *Praat* scripting section on the *Praat* manual. A copy of this script (and of the one that draws the spectrograms[9]) is given outside this written document. In what follows, the settings of the main analysis are detailed.

9.5.1. Formant settings

The values of the formants are drawn automatically through the *Praat* software. Specific settings are required to extract these values from the speech files. The parameters for extracting the formant values are assigned according to the indications provided by the *Praat* manual.

The maximum number of formants extracted per frame is five. The maximum value per formant (formant ceiling) is different according to gender: 5,500 Hz for a female voice and 5,000 Hz for a male one. (Go back to some of the spectrograms in the examples so far provided and look at the maximum values up in the frequency-vertical axis for both males and females subjects. You will see this difference!) The analysis window is set to 0.025 seconds, therefore the values of the frequencies are drawn each 25 ms of sound. The pre-emphasis value is set from 30 Hz.

9.5.2. Spectrogram settings

Spectrogram settings are concerned with the realisation of the spectrogram outlook. Also here, all spectrogram settings are assigned according to the indications provided by the *Praat* manual.

The window length is set at 0.005 seconds, that is *Praat* will analyse the 0.0025 seconds of sounds before and 0.0025 seconds after the centre of the window. The maximum frequency is set according to gender, 5,500 Hz for female speakers and 5,000 Hz for males. The time step is set at 0.002 seconds, the frequency step at 20 Hz. The window shape is Gaussian, pointed by the manual as the best choice among all others.

9.5.3. Harmonicity settings

In order to compute the value of the HNR value of the sounds we need to create the Harmonicity object in *Praat*. In this process too, the default values assigned by *Praat* are kept for the computation of the values. However, one exception is made in this domain and it will be explained later.

The preferred method according to the manual is cross-correlation. The duration of the frame on which the computation is done, time step, is the standard value of 0.01 sec. The minimum pitch is set according to gender, 100 Hz for female speakers and 75 Hz for males. The silence threshold is set

9 This is basically the same thing as the previous one. The core section is identical, that is the analysis for the classification of sounds. The difference is that here only a word of a speaker is analysed, not her whole repertoire. Secondly, this second script does not render a list in the *Praat* info window but a spectrogram in the Praat picture window. Actual examples of the results of this script are the spectrograms displayed from section § 8. to section § 9.

at 0.1, it means that those frames not containing amplitudes above this value are assumed to be silent. As said, the only refinement to the default parameters, is the reduced number of periods per window which has been lowered from 4.5 to 1.0. This betterment prevents *Praat* from stopping the running of the scripts because in some cases the widow length was longer than the sound (cases of short segments, around or below 35/40 ms duration).

9.5.4. Intensity settings

The intensity of the sound is not an essential hallmark for the classification of the sounds of this research. However, in some cases it helps in spotting whether a sound is a trill or a tap. If you see the spectrogram in fig. 8.5, the cyan line of the intensity contour shows an evident up-and-down pattern. This is due to the alternation of the closed and opened phase occurring when the tongue touches a fixed organ more times in a typical trill production. It means, basically, that during the closed phase the pressure of the sound lowers dramatically wheres in the opened phase the opposite occurs. In some cases, the sound was so unclear and blurred, that, although at least one initial white band was present, the rest were difficult to judge as other blurred vibrations of an unclearly shaped trill sound or the typical friction-noise of a released tap. The pattern of the intensity solved this type of problem.

Settings are: 100 Hz as the minimum pitch; the time step is set to 0.0, this means that the time step is one quarter of the effective window length. The third and last setting of the intensity contour allows *Praat* to subtract from the pressure of the recorded sound the constant air pressure that many devices, such as the microphone employed for the recording session, might have added. This drawback results in a non-zero value of the intensity in the sound wave even in silent phases of the recordings. *Praat* computes its mean and subtracts it from the intensity of the actual recorded speech. The setting is thus set to *yes* [subtract mean pressure].

9.5.5. Voice settings

The presence of the F_0 in a segment ascertains whether the sound is voiced or not. In the above spectrograms, the blue line represents the pitch. To draw the key information about this important feature, pitch settings are assigned according to the indications provided by the *Praat* manual.

Therefore, the range for F_0 is set at 75 Hz to 300 Hz for male speakers and at 100 Hz to 500Hz for female speakers. During the phase of segmentation it was clear that some /r/ sounds were neither completely voiced nor completely unvoiced. To determine whether a sound is [+voice] or [-voice] a rather complex plan has been designed into the script.

Taking the F_0 mean value from the starting point of the sound to its end, as it was done for formant values, was problematic because for those parts lacking of F_0 (that is, when the pitch contour stops and the blue line disappears) *Praat* gives *-undefined-* values. This parameter is misleading, and above all, we are not interested in the actual values of the pitch but, rather, on its mere presence throughout the /r/ segment.

In order to judge it as voiced or unvoiced the segment is thus split into nine sections. Each sections measures 2/16 of the sound (each sixteenth is one *step*). The centre of the first sound frame matches with the begin of the /r/ sound, in the *Praat* script this is called *beginborder*. The total length of the frame is calculated by adding one sixteenth to the right and one sixteenth to the left of the centre. So that the first section begins one step earlier in the beginning of the /r/ segment and ends one step after. The centre of the second section is at *beginborder + (step * 2)*. This procedure is repeated until the end of the /r/ segment, the *endborder*, is reached as being the ninth centre.

The segments presenting at least seven sections out of nine with definite F_0 means (that is, they do not give *undefined* outcomes) regardless of the value in Hz, are considered as *voiced*. Sounds showing fewer than four voiced sections are considered *unvoiced*. Cases in between (4 to 7 voiced sections), are too weak to be considered voiced but at the same time too strong to be considered unvoiced. Accordingly, these cases are labelled as *0voice*.

9.6. Summary

In summary, the analysis of the sounds of the sample is carried out by a script made through *Praat* 5.3. It is a semi-automatic program, because some sounds are previously detected by eye i.e. trills, taps, epenthesis insertions and zero. The sounds that are left out by the preliminary manual analysis are classified through the constraints set into the *Praat* scripts. This program takes into account the acoustic of the sounds and their analysis can be summarised through the following statements:

- if tier 2 is not empty, the analysis will classify it according to the cue itself (trill, tap, etc.)
- if its F_3 mean value is below 2,000 then the sound is an approximant
- if it is not one of the above and the HNR value is below 8 dB the the sound is a fricative
- if the HNR value is above 8 dB then the sound is a vowel.

10. Results

10.1. Tables and graphs: main findings

SPEAKER	Age Group	Gender	Total % of realisations		
			Appx	Rolled (trills + taps)	NON-Appx
F22ABOO22	Y	F	69,23	0,00	30,77
F24CMBOOO02	Y	F	90,95	0,00	9,05
F21HBAA21	Y	F	39,52	0,00	60,48
F22IMAA22	Y	F	84,69	0,00	15,31
F22JBAA22	Y	F	63,33	0,00	36,67
F21LBAA21	Y	F	43,41	0,00	56,59
F24MMAA24	Y	F	92,34	0,00	7,66
F24MMOA06	Y	F	90,34	3,86	9,66
F24MBAA24	Y	F	41,35	6,25	58,65
F26NMAA26	Y	F	4,78	25,36	95,22
F23SMAA23	Y	F	14,01	16,91	85,99
F23SMOA06	Y	F	90,48	0,48	9,52
F22TBAA21	Y	F	42,51	0,00	57,49
M18AVMBOAA18	Y	M	14,78	0,49	85,22
M27DMAA25	Y	M	79,90	8,61	20,10
M24DBOA03	Y	M	3,86	65,22	96,14
M20GBAA20	Y	M	73,91	0,00	26,09
M18JHAVOAA18	Y	M	83,41	0,00	16,59
M27JMOA19	Y	M	66,50	2,91	33,50
M22JMAA22	Y	M	88,57	0,48	11,43
M25KMOA10	Y	M	44,71	0,00	55,29
M20LBOA02	Y	M	91,87	1,91	8,13
M22TBOA04	Y	M	2,90	65,70	97,10
F61AHBOAO61	O	F	0,48	49,52	99,52
F56EMAA56	O	F	29,61	32,52	70,39
F60FHBOAA35	O	F	5,31	57,00	94,69
F59MHBOAA54	O	F	2,40	15,38	97,60
F66MMOA46	O	F	0,48	12,02	99,52
F60OPHDOA34	O	F	0,48	62,50	99,52
F64THBOAA64	O	F	11,00	32,54	89,00
F51TVWOAA51	O	F	13,94	0,00	86,06
F56YMAA56	O	F	21,05	36,36	78,95
M60BMAA47	O	M	21,43	7,14	78,57
M62GHBOOA51	O	M	1,43	1,43	98,57
M59HPHDOA31	O	M	85,85	0,98	14,15
M50JMBOAA50	O	M	0,00	28,43	100,00
M70MMAA70	O	M	12,86	3,81	87,14
M64PMOA30	O	M	42,86	0,00	57,14
M68PHBOAA28	O	M	21,15	5,77	78,85
M57PHBOOA34	O	M	34,80	42,16	65,20

Table. 10.1. Percentages of the total realisations of approximants, rolled and non approximant for the whole sample. (The third column 'NONApp' adds up the rolled and the other non-approximant-like articulations).

SPEAKER	Approximant				Trill				Tap		ə inser.		Fricative		schwa	zero	Vowel	Total
	ɹ	ɹ̥	ɻ	ɻ̊	r	r̥	ʀ	ʀ̥	ɾ	ɾ̥	ɾə	ɾ̥ə	F	F̥	ə	∅		
F22A	23	0	120	1	0	0	0	0	0	0	0	0	2	4	26	1	31	208
F24C	51	0	138	2	0	0	0	0	0	0	0	0	0	0	16	0	3	210
F21H	3	0	80	0	0	0	0	0	0	0	0	0	2	0	38	0	87	210
F22I	16	0	157	4	0	0	0	0	0	0	0	0	0	0	19	0	13	209
F22J	21	1	111	0	0	0	0	0	0	0	0	0	0	1	35	0	41	210
F21L	24	4	56	5	0	0	0	0	0	0	0	0	16	3	54	0	43	205
F24M	41	1	151	0	0	0	0	0	0	0	0	0	1	0	10	0	5	209
F24M	26	3	155	3	3	0	4	0	1	0	0	0	1	0	7	0	4	207
F24M	17	1	67	1	0	3	1	1	1	7	2	0	10	6	17	2	72	208
F26N	2	1	7	0	5	2	0	0	16	30	0	0	14	57	23	3	49	209
F23S	7	2	17	3	0	0	0	0	15	20	10	0	24	42	15	0	52	207
F23S	35	0	155	0	0	0	0	0	0	1	0	0	1	1	9	3	5	210
F22T	17	4	62	5	0	0	0	0	0	0	5	0	29	8	33	0	44	207
M18A	29	0	1	0	0	0	0	0	1	0	0	0	12	25	56	12	67	203
M27D	10	12	95	50	0	3	0	0	2	13	0	0	9	11	1	0	3	209
M24D	1	0	7	0	1	2	14	7	28	83	20	4	13	15	3	3	6	207
M20G	39	0	114	0	0	0	0	0	0	0	5	0	2	2	16	0	29	207
M18J	78	2	89	2	0	0	0	0	0	0	4	0	5	2	6	4	13	205
M27J	22	1	111	3	0	0	0	0	1	5	2	3	13	7	14	0	24	206
M22J	18	1	163	4	0	0	0	0	0	1	1	0	7	3	6	1	5	210
M25K	7	0	86	0	0	0	0	0	0	0	9	0	13	78	1	0	14	208
M20L	91	0	101	0	2	2	0	0	0	0	0	0	1	2	4	1	5	209
M22T	4	0	2	0	1	5	3	12	2	113	10	10	7	24	4	0	10	207
F61A	0	0	1	0	4	6	10	11	11	61	4	0	18	63	8	1	10	208
F56E	0	0	61	0	4	3	59	0	1	0	14	0	1	2	26	0	35	206
F60F	1	0	10	0	42	30	11	0	17	18	13	5	16	18	7	3	16	207
F59M	1	0	4	0	8	1	8	0	7	8	21	2	7	6	60	1	74	208
F66M	1	0	0	0	5	1	10	2	4	3	25	1	15	30	23	1	87	208
F60O	0	0	1	0	7	36	0	0	6	81	12	10	2	24	14	2	13	208
F64T	15	0	8	0	6	12	1	0	7	42	8	2	12	27	32	6	31	209
F51T	5	0	24	0	0	0	0	0	0	0	0	0	0	1	48	0	130	208
F56Y	5	0	39	0	1	6	7	3	4	55	17	0	6	23	25	0	18	209
M60B	25	0	20	0	2	2	7	0	0	4	22	5	19	94	5	0	5	210
M62G	3	0	0	0	0	0	2	1	0	0	12	0	7	183	0	0	2	210
M59H	8	3	145	20	0	1	1	0	0	0	8	0	6	4	4	1	4	205
M50J	0	0	0	0	1	0	3	0	18	36	23	2	33	62	2	15	9	204
M70M	22	1	3	1	0	0	0	0	2	6	16	3	53	13	29	1	60	210
M64P	37	0	53	0	0	0	0	0	0	0	9	0	4	1	40	4	62	210
M68P	19	0	25	0	2	0	10	0	0	0	2	0	9	3	29	2	107	208
M57P	21	0	50	0	10	8	13	7	17	31	17	6	8	2	9	1	4	204

Table. 10.2. Frequencies of the sounds produced by tall the speakers of the sample. Sounds are classified as voiced or unvoiced approximant through the automatic *Praat* extraction script. Trills, alveolar vs. uvular and voiced vs. unvoiced, are detected by eye. So are taps, schwa insertions and zero. Fricatives (remember that the [F] stands for any fricative) and vowels are picked out through the script.

Once the Praat script was finalised the stage of actual analysis took place. The analyser classifying script was run for each speaker. The results were large tables for each participant containing the list of all the words of the sample, the phonological contexts of the /r/ segments, the mean values for F_1, F_2 and F_3, and most importantly, the classification of the /r/ segments through the appropriate IPA symbol.

These outcomes were given in tables through the *Praat* info window. They were copied and pasted into electronic sheets in *Numbers '09* so that for each speaker one electronic sheet was created. For each speaker table the results were analysed as follows: first, all the items were rearranged and put in order according to their phonological contexts (since the words were completely scrambled throughout the test); second, each occurrence of a sound was counted and summed up. If a sound was made-up by two segments, only the first one was considered. Third, the outcomes were turned into percentages. When all the summary tables were achieved for each speaker, a bigger, unique electronic sheet table was created containing all the relevant information needed for developing graphs and statistical analysis.

Table 10.1 shows the percentages of rhotic approximants ([ɹ] and [ɻ]), rolled /r/s ([r] and [ɾ]) and others (fricatives, vowels or zero) for each speaker e.g. F24CMBOOO02, female of the young group, used an approximant sound 90,95% of the times out of 210 items[10]. Table 10.2, in turn, is more detailed than the previous one because it presents all the occurrences of each sound for each speaker. Both tables were created through the final electronic sheet table.

Among old speakers the rhotic approximant is rather infrequent. Surprisingly, it is more spread in male speakers (27.61%) than in females (9.42%). As far as rolled sounds are concerned, there are remarkable high percentages among female speakers.

Both young groups respect the expectations as well, but some surprises arise. In general, both genders make a larger use of approximants with only five speakers out of twenty-three presenting in their repertoires percentages below 40% of approximants in their repertoires. Rolled /r/s seem to disappear in the young female group because this sound is utterly missing from the speech of eight out of thirteen speakers and only two have a percentage above 6%. Male speakers show a higher degree of variation compared to their female peers. Rolled /r/s are the main pronunciation for two speakers out of ten, but generally approximants are quite widespread.

In order to demonstrate the difference between the four groups, four bar graphs have been generated fig.10.3.

10 The total number of the items in the reading material is 210. Considering that some speakers missed or misread some words during the test, this number in many cases lowers down of some units. The 'worst' reader gave 203 analysable sounds.

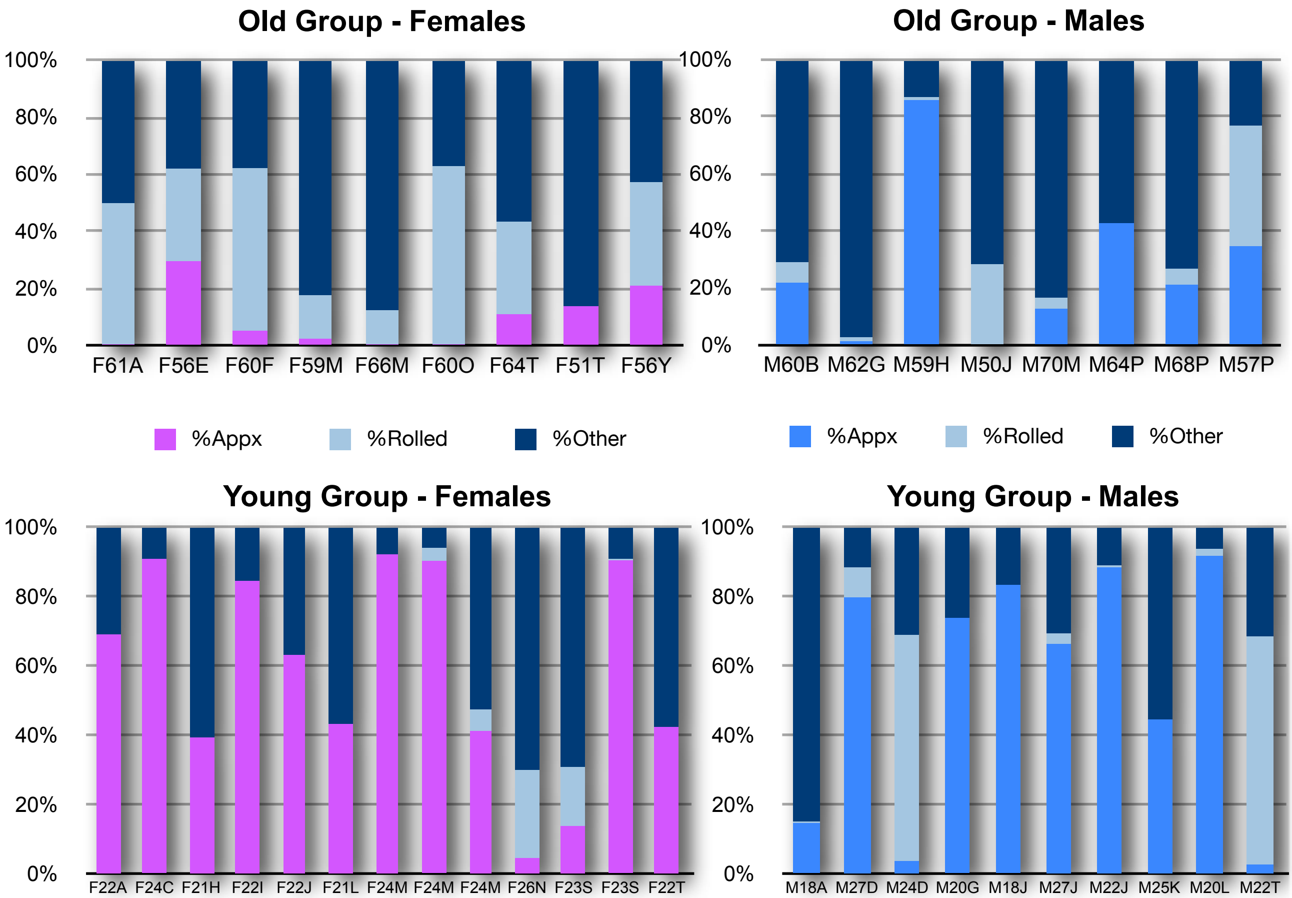

Fig. 10.3. Graph bars representing the total percentages of approximant, rolled /r/ and other pronunciations divided into the four groups: old females, old males, young females and young males.

10.2. Statistical analyses

The differences seen in the tables and in the bar graphs above are not enough to prove anything. To test the hypothesis of diachronic change and to discover more patterns of change throughout our sample, several statistical tests have been carried out. The questions we want to answer are the following:

a - Do young speakers speak differently from old speakers?

b - Do female speakers speak differently from male speakers?

c - Do different phonological contexts affect the pronunciation of the speakers?

d - How are the different sounds distributed across age groups?

e - Is there any correlation between age and the F_3 mean?

f - Is there any correlation between the new sound and the epenthetic schwa?

10.2.1.Differences across age and gender

This sections tries to answer to both questions *a* and *b*. To test the total means of approximant, rolled, and non-approximant sounds three different analyses of variance were carried out. The between factors of the analyses are age and gender. Age always appears significant for all the three sounds. The discrepancy in the usage of approximants in the young group (M = 57.28%, SD = 32.05%) compared to the old group (M = 17.98%, SD = 21.9%) is statistically significant $F(1, 36) = 17.97$, $p = 0.00015$. Rolled sounds are much more common in the old group (M = 22.80%, SD = 21.33%) than the young group (M = 8.62%, SD = 18.99%) with $F(1, 36) = 4.58$, $p = 0.039$. Finally the use of sounds that are not an approximant (rolled plus any non-approximant) is more frequent in the old group (M = 82.05%, SD = 21.92%) compared the young group (M = 42.72%, SD = 32.05%) with $F(1, 36) = 17.97$, $p = 0.00015$.

As far as the gender factor, the outcomes of the statistical analyses always show not significant differences in the usage of the three sounds. This seems to point that gender does not play a role in the sound change under investigation. This is rather disappointing since the hypothesis that 'women lead the change' (section § 5.3.) was disproved.

Nonetheless, in one case the interaction between gender and age is statistically significant. The usage of rolled sounds seems to be dependant on the interaction between age and gender $F(1, 36) = 7.25$, $p = 0.011$. This has to be related to the high percentages of rolled sounds for the old female group (M = 33.09% SD = 21.13%) compared with the other three groups (OM: M = 11.21% SD = 15.50%; YF: M = 4.07% SD = 8.00%; YM: M = 14.53% SD = 26.97%).

From the results shown above we have been able to prove that the difference between the two age groups is statistically significant. Accordingly, we can infer that a diachronic change, leading a rolled-like, non-approximant sound in the /Vr/ context towards an approximant sound, is under way in the Amsterdam Dutch.

10.2.2.Difference across phonological contexts

To test whether the different phonological contexts may affect speakers' pronunciations, several analysis of variance tests have been run. The following table collects the results of all the analysis of variance carried out for this stage of the statistical analysis. Age is expected to be consistently significant. At the same time, we are curious to see whether other variables (stress, vowel length, etc) may give us further insights on the usages of the newer and older sounds.

As we can see from the results, the two-factor repeated measures analysis of variance always shows a highly significant effect for the use of approximants between the two age groups. This, again, proves that the diachronic phonetic change theory is soundly based. Gender does not show significant different effect in any test. The analysis of variance results, however, show that the interaction between gender and the length of the preceding vowel is significant, $F(1, 36) = 11.108$, $p = .002$. This tells us that the use of an approximant rhotic sound is more likely to occur when the speaker is a male and, at the same time, the preceding vowel is a long one. (Male speakers in LVappx total mean = 44.50% SD = 35.08% vs. M = 36.67 SD = 35.14% for female speakers).

Test	Factors				Results	Significant?	
Appx	Age & Gender	vs.	Final context	The occurences of the appxs in the context V/r/# against the occurences of the appxs in V/r/C#	F(1, 36) = 0.247, p > .05	Final context	NO
					F(1, 36) = 17.90, p = .00015	Age	YES!
					F(1, 36) = 0.625, p > .05	Gender	NO
					F(1, 36) = 1.457, p > .05	Age*Gender	NO
					F(1, 36) = 2.652, p > .05	Age*Final context	NO
					F(1, 36) = 0.839, p > .05	Gender*Final context	NO
					F(1, 36) = 0.772, p > .05	Age*Gender*Final context	NO
Appx	Age & Gender	vs.	Vowel length	The occurences of the appxs in the context VV/r/ against the occurences of the appxs in V/r/	F(1, 36) = 0.064, p > .05	Vowel length	NO
					F(1, 36) = 18.17, p = .00014	Age	YES!
					F(1, 36) = 0.510, p > .05	Gender	NO
					F(1, 36) = 1.471, p > .05	Age*Gender	NO
					F(1, 36) = 0.149, p > .05	Age*Vowel length	NO
					F(1, 36) = 11.108, p = .002	Gender*Vowel length	YES!
					F(1, 36) = 0.084, p > .05	Age*Gender*Vowel length	NO
Appx	Age & Gender	vs.	Vowel position	The occurences of the appxs in the FV/r/ against the occurences of the appxs in CV/r/ and BV/r/	F(2, 72) = 16.523, p = 1.24*10-	Vowel position	YES!
					F(1, 36) = 19.546, p = 8.67*10-	Age	YES!
					F(1, 36) = 0.560, p > .05	Gender	NO
					F(1, 36) = 1.400, p > .05	Age*Gender	NO
					F(2, 72) = 0.054, p > .05	Age*Vowel position	NO
					F(2, 72) = 3.881, p > .05	Gender*Vowel position	NO
					F(2, 72) = 6.177, p = .018	Age*Gender*Vowel positio	YES!
Appx	Age & Gender	vs.	Vowel stress	The occurences of the appxs in the 'V/r/ against the occurences of the appxs in V/r/	F(1, 36) = 6.288, p = .017	Vowel stress	YES!
					F(1, 36) = 17.798, p = 0.00016	Age	YES!
					F(1, 36) = 0.647, p > .05	Gender	NO
					F(1, 36) = 1.418, p > .05	Age*stress	NO
					F(1, 36) = 0.638, p > .05	Age*Vowel stress	NO
					F(1, 36) = 2.163, p > .05	Gender*Vowel stress	NO
					F(1, 36) = 1.774, p > .05	Age*Gender*Vowel stress	NO

Table 10.4. Details of each analysis of variance with repeated measures for the % of rhotic approximants.

The phonological contexts that most affect the use of the approximants are two: the position of the vowel within the vocalic quadrilateral and the presence of the stress in the preceding vowel. Let's see these points in detail. The position of the vowel is also significant for the following sound. The analysis of variance showed a significant difference on the position of the preceding vowels, $F(2, 72) = 16.523$, $p = 1.24*10^{-6}$. Considering the mean percentages of approximants per vowel position (FV: M = 36,16% SD = 35.09%, CV: M = 35,39% SD = 34.04%, BV: M = 48,65% SD = 35.30%), we can draw that back vowels are more likely to be followed by an approximant than central and front vowels. Moreover, vowel position interacting with gender and age also gives a significant result $F(2, 72) = 6.177$, $p = .018$.

Stress also seems to affect the realisation of the following sound. Unstressed vowels are more likely to be followed by an approximant rather than stressed ones (ˇV: M = 42.51% SD = 34.31%, ˈV: M = 38.94% SD = 34.69%).

The following table resumes, like the previous one, the statistical analysis outcomes obtained through analysis of variance with the same factors but through the data of another dependant variable, namely the percentages of rolled sounds.

Test	Factors		Results	Significant?	
Rolled	Age & Gender vs. Final context	The occurences of the rolled sounds in the context V/r/# against the occurences of the rolled sounds in V/r/C#	F(1, 36) = 15.321, p = .00039	Final context	YES!
			F(1, 36) = 4.523, p = .04	Age	YES!
			F(1, 36) = 0.893, p > .05	Gender	NO
			F(1, 36) = 7.290, p = .01	Age*Gender	YES!
			F(1, 36) = 3.442, p > .05	Age*Final context	NO
			F(1, 36) = 0.576, p > .05	Gender*Final context	NO
			F(1, 36) = 0.856, p > .05	Age*Gender*Final context	NO
Rolled	Age & Gender vs. Vowel length	The occurences of the rolled sounds in the context VV/r/ against the occurences of the rolled sounds in V/r/	F(1, 36) = 5.034, p = .031	Vowel length	YES!
			F(1, 36) = 4.835, p = .034	Age	YES!
			F(1, 36) = 0.986, p > .05	Gender	NO
			F(1, 36) = 7.384, p = .01	Age*Gender	YES!
			F(1, 36) = 4.037, p > .05	Age*Vowel length	NO
			F(1, 36) = 2.974, p > .05	Gender*Vowel length	NO
			F(1, 36) = 0.078, p > .05	Age*Gender*Vowel length	NO
Rolled	Age & Gender vs. Vowel position	The occurences of the rolled sounds in the FV/r/ against the occurences of the rolled sounds in CV/r/ and BV/r/	F(2, 72) = 2.849, p > .05	Vowel position	NO
			F(1, 36) = 4.762, p = .036	Age	YES!
			F(1, 36) = 0.429, p > .05	Gender	NO
			F(1, 36) = 6.292, p =.017	Age*Gender	YES!
			F(2, 72) = 0.656, p > .05	Age*Vowel position	NO
			F(2, 72) = 0.967, p > .05	Gender*Vowel position	NO
			F(2, 72) = 0.096, p > .05	Age*Gender*Vowel positio	NO
Rolled	Age & Gender vs. Vowel stress	The occurences of the rolled sounds in the 'V/r/ against the occurences of the rolled sounds in V/r/	F(1, 36) = 31.023, p = 2.61*10-	Vowel stress	YES!
			F(1, 36) = 4.500, p = 0.041	Age	YES!
			F(1, 36) = 0.837, p > .05	Gender	NO
			F(1, 36) = 7.122, p = .011	Age*Gender	YES!
			F(1, 36) = 4.684, p = .037	Age*Vowel stress	YES!
			F(1, 36) = 8.066, p = .007	Gender*Vowel stress	YES!
			F(1, 36) = 8.273, p = .007	Age*Gender*Vowel stress	YES!

Table 10.5. Details of each analysis of variance with repeated measures for the % of rolled sounds.

Once again, age is always significantly different between the subject groups. However, the *F* ratios are lower compared to those obtained in the approximant analysis, therefore the chance to commit Type I error is higher. The difference is still significant but not as much as we would have expected in a first place. This might be due to the old group figures (especially males), which showed a rather low percentage of rolled sounds but a very high percentage of non-approximant and non-rolled sounds. This made the difference between the two subject groups significant yet not enough to exceed the .01 alpha level. Again, gender itself is never significant for each of the tests carried out. However, in each context the interaction between age and gender is significant. This is due to significant differences in female usages of rolled sounds: old women use more rolled sounds than young whereas no such difference is seen between males.

The final context is highly significant, that is rolled sounds are influenced by the coda of the syllable: when /r/ is followed by another consonant trills and taps are less likely to be produced (M =

9.91% SD = 17.94%) than in V/r/# position (M = 19.02% SD = 25.91%). As we anticipated, the interactions between age and gender are always significant for rolled sounds. Here, is significant at F (1, 36) = 7.290, p = .01. This result becomes clear when we look at the means of the rolled sounds split up by gender and age. We discover that there is a large difference between old (M = 41.00% SD = 24.21%) and young women (M = 6.32% SD = 12.04%) in the /r/_# domain and (M = 23.47% SD = 22.36%) and (M = 1.61% SD = 3.95%) in the /r/_#C. The interaction of the gender-variable with age therefore becomes significant.

The length of the vowel, unlike with appx sounds, is significant: $F(1, 36) = 5.034$, $p = .031$. The difference among the two groups are narrow: VV (M = 13.79% SD = 21.47%) and V (M = 15.77% SD = 20.86%) but the test showed a significant effect, that is short vowels are more likely to be followed by rolled sound than long vowels.

The position of the vowel does not affect the production of a rolled sound, whereas stress shows a significant effect: after a stressed vowel trills and taps are produced more often, (M = 16.90% SD = 22.81%) than after unstressed vowels (M = 11.93% SD = 19.24%). Interestingly enough, all the possible interactions in this domain give significant results.

10.2.3.Distribution across age

The table in fig. 10.6. is meant to give an outline of the overall realisations broken into ten different age groups.

The first four groups on the x axis, the young speakers (from 18-20 to 25-27), show a remarkable regularity. Rolled /r/s are rare and span from 0% to 15% of the total production of the postvocalic /r/s. Also the approximant /r/s show a good stability but, as expected, values are much higher and roughly cover the range from 50% to 65% of the total.

The old groups show an overall low use of approximants, from 0% to around 30% in one age group (56-57). Trills and taps are very common although there is an unpredicted fall of the figure for the older groups. The burgundy line instead reaches very high percentages up to 80% in two occasions for the old groups. This last figure is possibly due to the heavy use of fricatives and/or vowels.

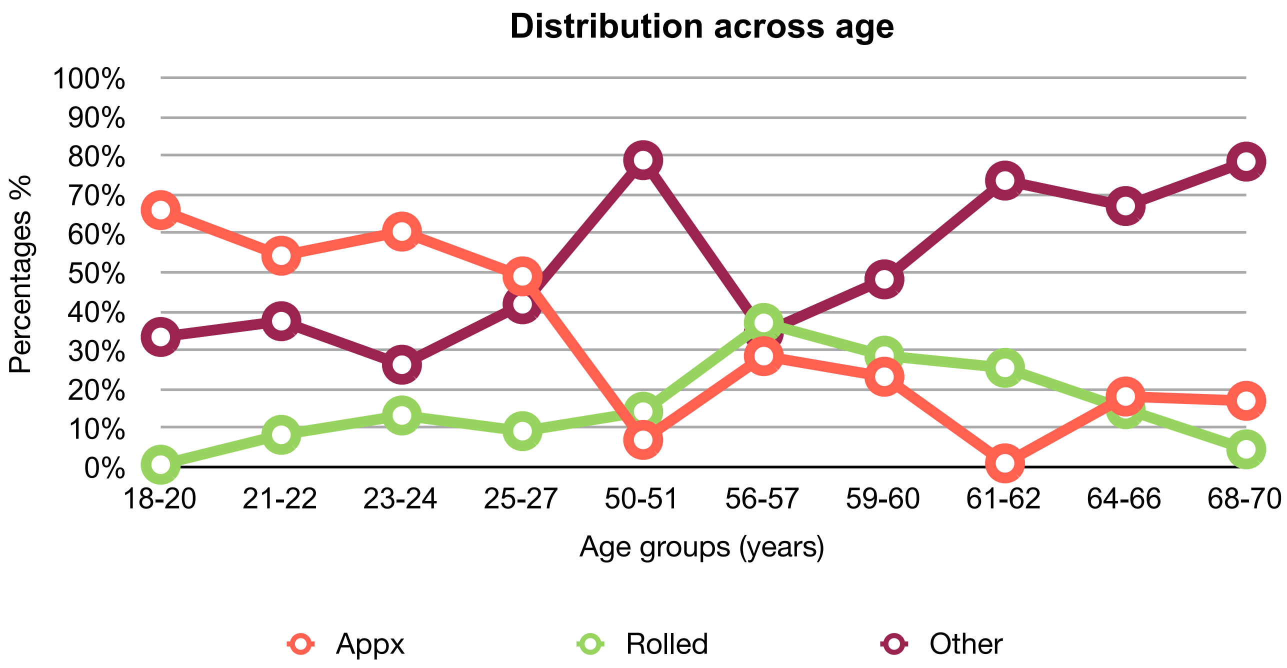

Fig. 10.6. Diagram representing the distribution across smaller age groups of the realisations of approximant and rolled.

Three major remarks must be highlighted. Firstly, the line representing the rolled /r/ begins to fall instead of keeping its rise. This was not expected. However, we previously saw that the rolled sounds were quite rare compared to our expectations and that non-approximant sounds were prominent through a more widespread usage of fricatives and vowels.

Secondly, the use of approximants seems to respect our expectations throughout the diagram. However, the last two groups of the old sample show an unexpected rise. Tentative explanations on this point will be provided in section § 10.2.5.

Thirdly, consistent with the hypothesis of the ongoing change, the variability is much higher in the old group, whereas for the young group diagram lines never cross between each other. Therefore, a stage of normalisation seems nowadays to take place. This is the core meaning of change: from a stage of variability, in which trills, taps, and others are clearly the bigger part, a new more regular stage is under way in which rolling seems to be disappearing and approximant rising in frequency. Think the above graph as it would move slowly rightwards, that is in the sense of the passing of time: we would see that the pattern shown by the young speaker figures remains more or less the same or even enhancing the current trend. The green line would remain fixed at 0% and the orange one should remains stable or even raise until, maybe, one day would reach 100%.

10.2.4. F_3 mean, correlation with age

The F_3 mean, that is the crucial boundary that discriminates approximant sounds from the rest, is seen here in deeper detail. Given this essential feature, it was interesting to compare the F_3 mean values throughout the whole sample. First of all, a sketch of the means for F_3 is provided in the following graph.

From the scatterplot below it is clear that F_3 values for both young males and females are rather close to 2,000 Hz, which is the agreed threshold for approximant sounds.

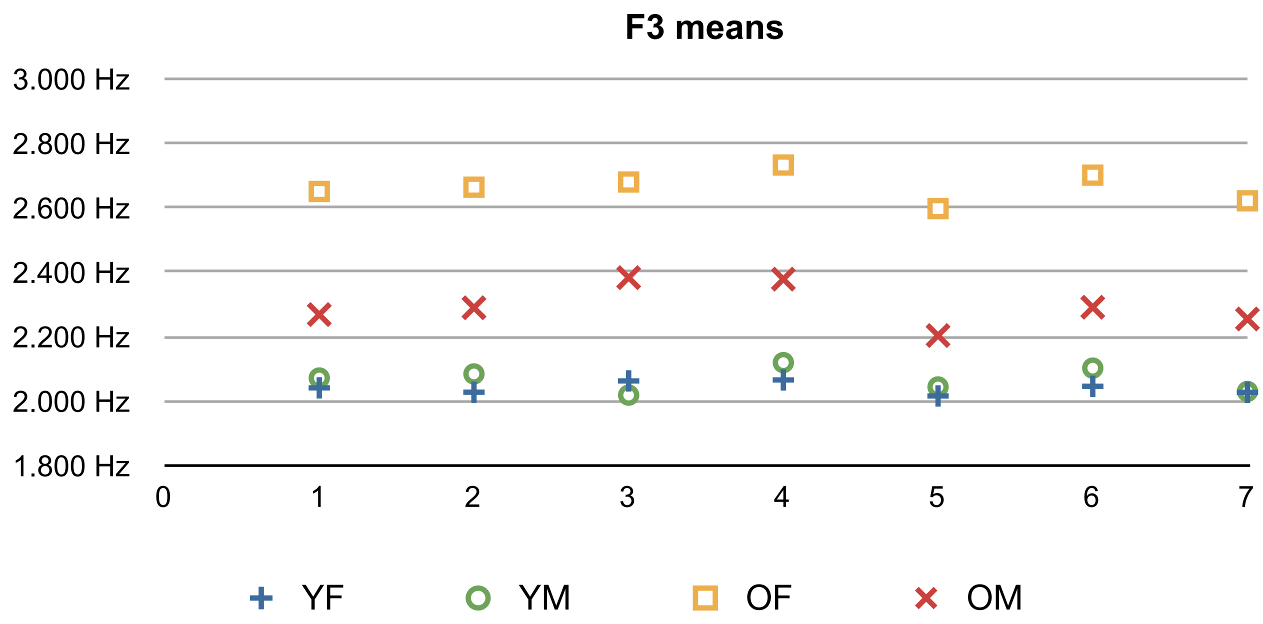

Fig. 10. 7 .Scatterplot of the F_3 means divided for age group and gender in different contexts. Y = young, O = Old, F = Female, M = male. 1 = after long vowel, 2 = after short vowel, 3 = after [ə], 4 = in /r/#, 5 = in /r/C#, 6 = after stressed vowel, 7 = after unstressed vowel.

These values suggest that within the young repertoires the prominent sound is represented by approximants since other sounds show higher values for that formant. However, since in all seven contexts the values for the young group exceed the threshold, we get new evidence that other pronunciations rather than approximants are still present (in fact at least eleven speakers out of twenty-three produce, in different frequencies, rolled, fricatives, vowels and zero sounds; see tables above).

As far as the old group, the means reflect a little use of approximants. Because of that, F_3 values go well beyond the 2,000 Hz threshold.

A major hint about the difference between the two age groups is provided by the comparison between F_3 means. Each F_3 value was averaged in all 42 contexts for each speaker. These means were averaged to find a total F_3 mean for each speaker of the sample. The twenty-three means of the young group were tested against the seventeen of the old's. The young group has an overall low F_3 mean (M = 2,053.74 Hz, SD = 310.22 Hz) than the old group (M = 2,478.59 Hz, SD = 372.00 Hz) with a significance of $t(38) = -3{,}93$, $p < .01$, two tailed.

The distribution of the F_3 means across age groups is provided below.

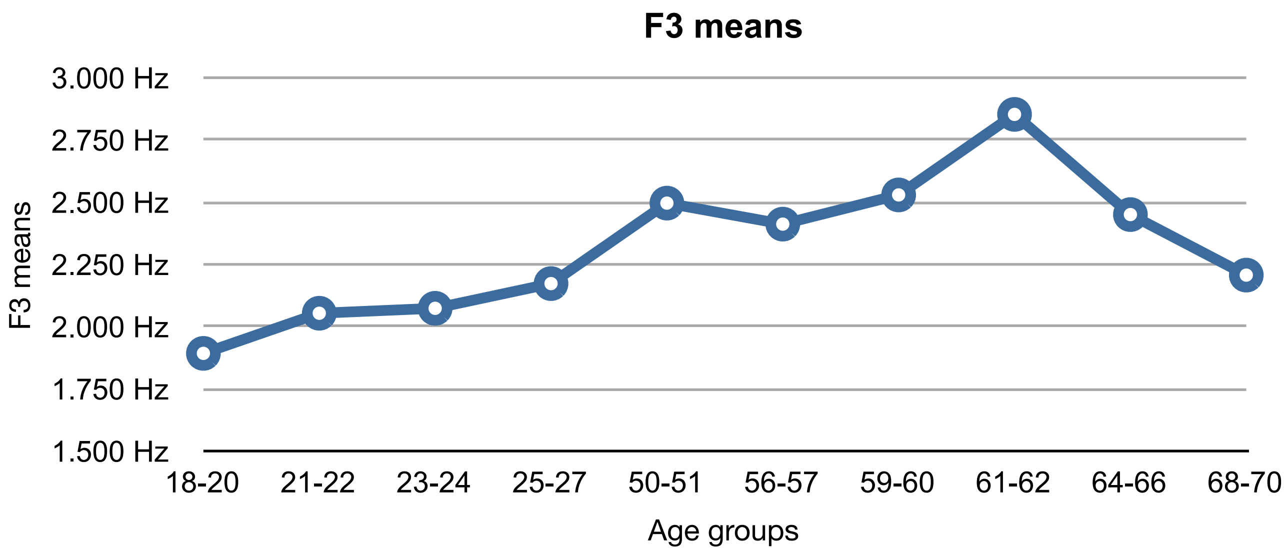

Fig. 10.8. Diagram representing the distribution across smaller age groups of the F_3 means.

According to this diagram, the younger you are the lower will be the overall F_3 mean value for each postvocalic /r/; that is, the younger you are the more frequently you will use an approximant. This utterly fits the hypothesis of the diachronic change. Furthermore, the diagram shows a constant rise of the F_3 mean value consequently to the rise of the age. Unexpectedly, there is a drop of the F_3 value for the two last age groups (64-66 and 68-70).

10.2.5. Unexpected results, tentative explanations

Having the informants record sheet checked, we see for the 64-66 group that a 64 year-old speaker actually spent half of his life in Den Haag (and also part of his childhood) bringing the mean of his group dramatically low. For the last group, made up only two speakers, a very low F_3 mean is brought by a 68 year-old speaker that spent more than half of his life outside Amsterdam. The oldest speaker also has a relatively low F_3 mean (around 2,300 Hz). Yet, he never left the city for any remarkable amount of time. So for a couple of cases which seem to point at a wrong selection of the speaker, the last one seems to be an unhappy coincidence. However, my guess is that this speaker slightly changed his pronunciation throughout his life especially because of his higher social status (assumed by his profession, notary).

Other points of discussion on the unexpected results provided by the diagram in fig. 10.8. may be arisen through the introduction of the notion of age-grading. One of the pitfalls of the apparent-time study is given by the phenomenon known as age-grading. Speakers do not change their speech habits when they have reached adulthood is the assumption on which this type of study stands. However, '[t]hat apparent-time reflects real time is a hypothesis, not an axiom or a theorem' (Chambers 2002:358). Age-grading takes place when 'members of a speech community alter their speech at some juncture of their lives in such a way as to bring it into conformity with adult norms' (Chambers 2002:358). Moreover, physical modifications of the phonatory apparatus may affect the way a (usually older) speaker talks. Labov (1994) mentions this phenomenon as a 'physical deterioration that interferes with their speech: loss of teeth, hoarseness of voice, and lax articulation' (Labov 1994:46).

Indeed, it cannot be drawn without a deeper analysis whether some of the old subjects were affected by either age-grading and/or physical deterioration. It must be reported however, that speaker M59HPHDOA31 had a denture. This may have affected to some extent his pronunciation and making the approximant rhotic sound much easier to produce compared to the trill sound (more common across his generation) in which the articulatory gesture is more demanding due to the tongue vibrating around the alveolar area of the mouth. He shows the highest percentage of approximant sounds in his category (Old males M = 85,85%).

Another comment may be made on speaker F51TVWOAA51. Data from her repertoire clearly stands out: she scored 0,00% as the production of rolled sounds. This data is far from the average of her fellow category participants (Old Females M = 37,23%). Taking into account that she is only 51 years old we can speculate that she has been able to adapt her speech to the new in-fashion linguistic norm. In so doing this speaker provides us an example of age-grading speech. Interestingly enough, she failed at producing *proper* approximant sounds (see section § 11.3.).

10.2.6.Schwa epenthesis and the production of approximant sounds

Throughout the experiment we came up with several cases of schwa epenthesis. It is a well-known feature in present day Dutch. However, as the rolled pronunciation in postvocalic position, the schwa epenthesis phenomenon seems nowadays receding. A negative correlation is therefore expected between the /ə/ insertion and the new approximant-like pronunciation of the /r/.

To test this hypothesis, each schwa epenthesis was counted. The suitable environments were / r/+velar [k, x, ɣ] or /r/+labial consonants [m, n, f, p]. In total, the appropriated items in the speech material for the epenthesis to be produced were twenty-four. Unexpectedly, a few [ə] insertions occurred when /r/ was in final position. One example is shown in the spectrogram in fig.10.10.

The statistical test of the correlation gave a significantly negative outcome. The data revealed that the use of the approximant /r/ and epenthetic [ə] were significantly related, $r = -.621$, $n = 40$, $p < .01$, two tails. This corroborates thus the hypothesis that the change that leads the postvocalic /r/ from rolled > approximant is somewhat tied with the deletion of this typical phonetic feature of Dutch.

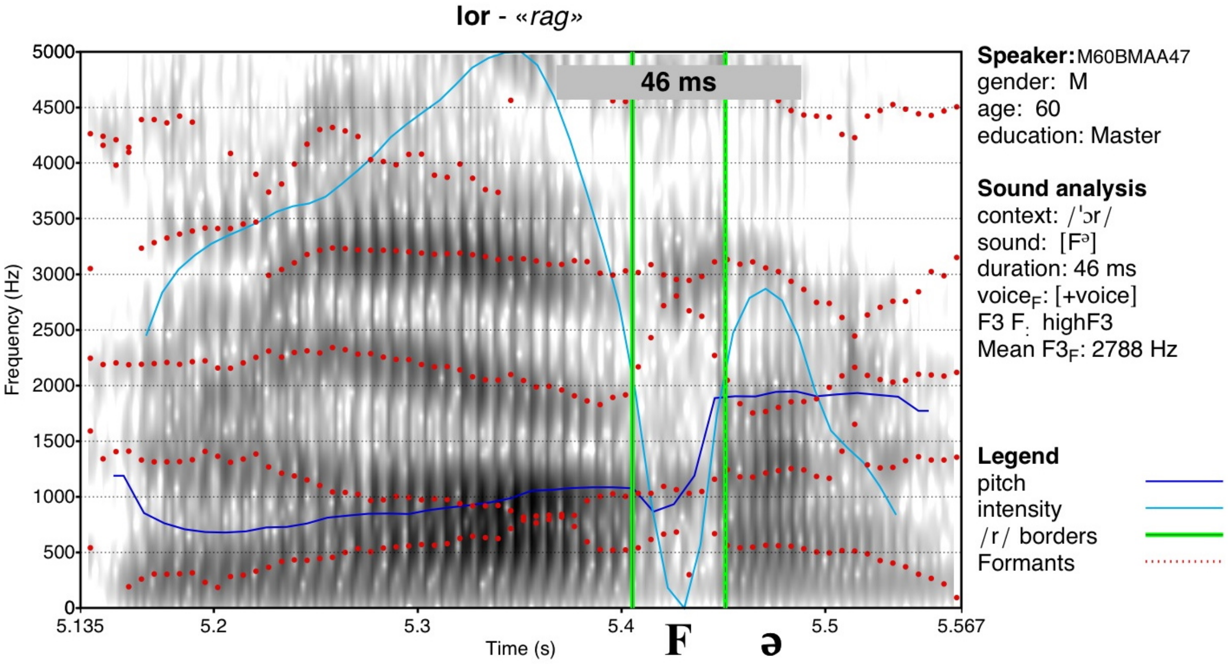

Fig. 10.10. Spectrogram of *lor* 'rag' (speaker M60BMAA47) performed through *Praat* v. 5.3. Although the word ends with an /r/ (/lɔr/) an epenthetic [ə] is added in final position. Cases such as this one are rare but, nonetheless, occurred throughout the experiment. Also, notice that the analysis gave a fricative sound. Nonetheless, this occurrence, like each case of epenthesis, is counted simply as [ɾə].

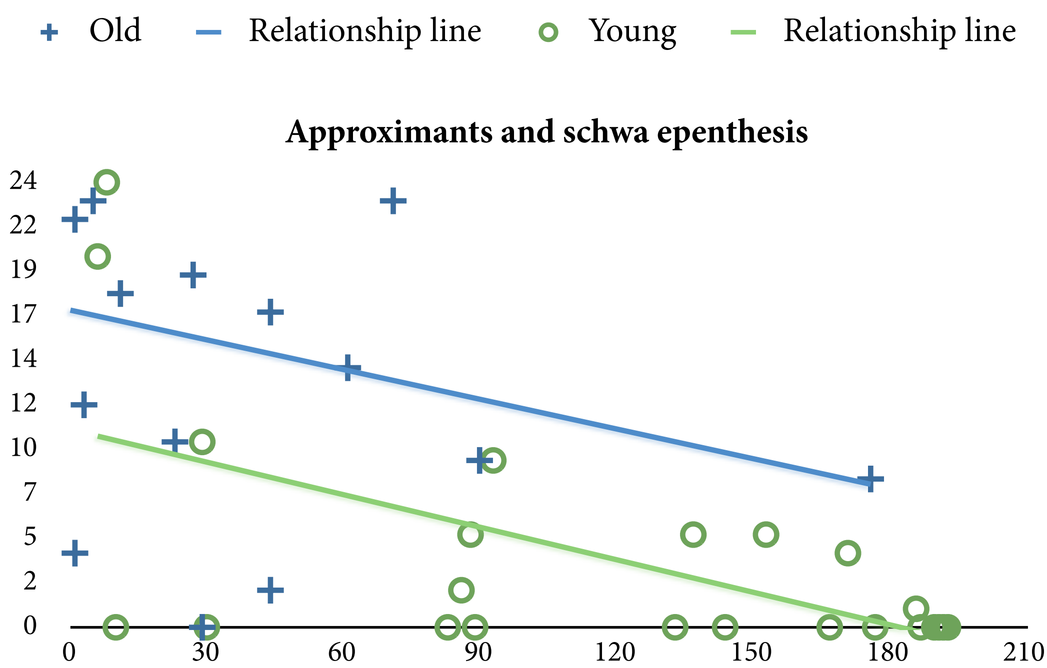

Fig. 10.11. Scatterplot of the raw scores of approximant /r/ (x) and epenthetic [ə] for each young speaker (green) and old speaker (blue). The tendency lines are almost parallel, this means that both groups behaves similarly, those speakers more prone to pronounce in the new way is, at the same time, giving up the schwa epenthesis.

11. Discussion

11.1. A brief look at acquisition and change

Among the subjects that took part in the test of this work there are two sons (one boy and one girl) and their mother. Thus, it is interesting to see their behaviour and whether they stay stick to their mother's accent or not. In other words, do the offspring talk like their mother or not? The answer that can be drawn here is a resounding *no*.

Figure 11.1. shows the pie charts for the percentages of these three speakers. Look first at speaker F60OPHDOA34. She has a 0% of approximant /r/ in her repertoire and more than 60% of rolled /r/ (the largest value of the entire work). One may expect to see a similar trend reflected in her sons' speech. Interestingly enough, it is the opposite that arises through the reading of the data: with some 80% of approximant pronunciations out of 210 items, her 27 year-old boy (M27DMAA25), places himself as one of the young male group speaker with the strongest tendency towards the approximant /r/. On the other hand, his sister (F22TBAA21) has a lower percentage of approximant in her repertoire (around 42%) but, surprisingly, there is no longer a trace of the rolled /r/ (0%).

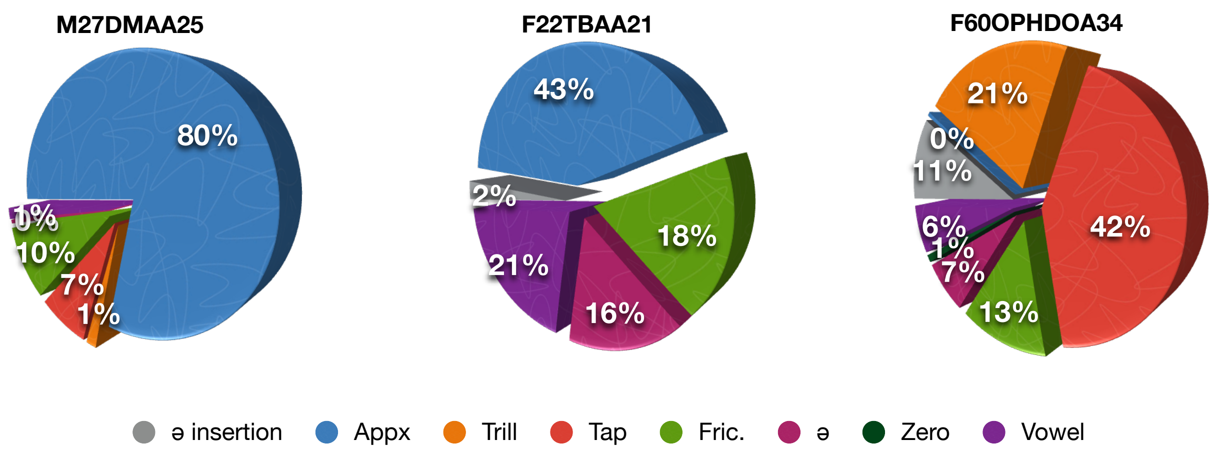

Fig. 11.1. Pie charts representing the distribution of the occurrences of the /r/s for three speakers. F60OPHDOA34 is the mother of the young speakers.

This gives evidence to the fact that speakers modify their language throughout their adolescenthood. If it is true that the first acquisition of the language hinges on the imitation of the parents' speech behaviours, it is also true that adult speakers are able and willing to change it according to new models of language available within the speech community.

Moreover, this brings new evidence to the current issue of acquisition of language and the role of speakers in language change: as far as this work is concerned, adult speakers are the only responsible. Here, the teaching-circle paradox seen in section § 5.1. is resolved. Despite parents are perfect teachers and children are perfect learners, language changes through innovations brought by adult speakers.

11.2. There is always more than the message itself

At the end of each recording sessions, speakers were challenged to guess what the real purpose of the test was about. I did not keep track of the answers but at least some of them (1 out of 4 roughly) succeeded in spotting that the test was concerning the *r*s. At that point I asked them what they had to say about their *r*s and about those of the other speakers. Apparently all of the speakers, regardless of their own type of *r*, acknowledged that the new variant is considered as a marker of a ‘posh’ accent.

Here arises an important lesson that must be learned: (adult) speakers are well aware that speech usually conveys an extra meaning to the mere semantics of a linguistic communication. In other words, the way we speak characterises us bringing additional information beyond the message contained in the chain of sounds we produce while speaking.

Extra cues may come from several sources: the tone of the voice, the use of marked synonyms, use of politeness (compare: "please, could I stay on my own for a while?" and "leave me alone!"), body language (a wink of the eye given while I'm speaking means, at least in Western cultures, ‘what I am saying is not true’), and pronunciation, which is, possibly, the strongest, the most effective, and common of all of the above.

Through that *extra* meaning coming out from the speech, we can guess from which area of our country a speaker comes from, and we may guess her position on the social scale. Interestingly enough, also the /r/ case in Dutch does not escape this rule ‘due to the high frequency of (r) in Dutch, in combination with its wide articulatory range, this phoneme is a potentially powerful cue for speaker identification’ (Smakman 2006:222).

11.2.1.The social evaluation of approximant sound in Amsterdam

So far, we established that the change [r] > [ɹ] is doubtless under way. We also discovered that those responsible for it are the adults and it has nothing to do with a child’s first acquisition. Therefore the question arises, what brought the emergence of the new sound? We believe that is due to a massive spreading of a trait already in place within Dutch dialects that is gaining acceptance through social evaluation.

If we agree that language changes, we should also agree with the metaphor that language is like fashion. As a consequence, therefore, language too may involve people’s taste or, better said, *style*. ‘Whatever “styles” are, in language or elsewhere, they are part of a *system of distinction*, in which a style contrasts with other possible styles, and on the social meaning signified by the style contrasts with other social meanings' (Irvine 2001:22) [her emphasis].

Considering that most of the informants agreed that in the D.o.A. the allophones [ɹ] and [ɹ̠] / _#, _C#, mark *poshness*, the extra meaning conveyed by these sounds is actually a sign of *distinction* that ascribes the speaker into the network of a new, younger, posh community in Amsterdam. It is startling that a tiny physical detail (the movements of the tongue that produces a low F_3) can be so powerfully meaningful!

However, the affiliation to a certain network is not systematic. As said, it is a matter of taste and style. I believe we can all agree that one is not what one is in actuality, but what one *thinks* of oneself. In other words, it is up to us to embrace a certain model of style based on our thoughts and tastes, that is a homeless person is not banned from using a posh discourse-style and, at the same time, the Prime minister's son is likely to speak more similarly to his peers rather than to his father.

The mere fact that *some* or *most* of the young people speak like *that* does not mean that each youngster will automatically adopt *that* same linguistic code. Rather, if they want to stay away from a certain network, they might signal a *contrast* (recall Irvine's quotation above) through their speech style. They therefore would avoid the most prototypical linguistic markers of that group. From the sample, two examples perfectly symbolise this sort of disobedience.

Informant M22TBOA04 admitted that after having gained the approximant /r/ in his language during his adolescence, he deliberately switched back to his previous one (because as a child he had learned the rolled /r/ from his parents, as he reported). He adduced that his choice was driven by his negative approach towards that accent. When adult, he became conscious of the extra meaning of that type of *r*. Another conscious refusal to the posh accent is brought by speaker F26NMAA26 who never renounced to her home's accent simply because she never really felt it as a necessary measure to enhance her social position or for whatever reason.

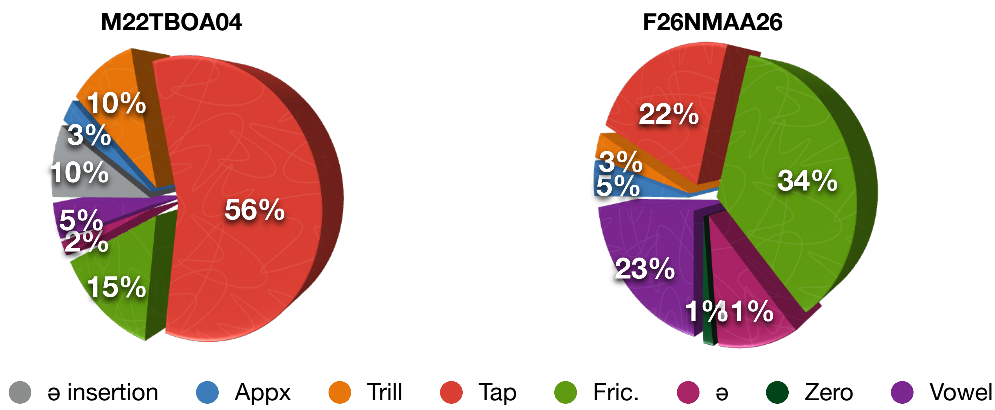

Fig.11.2. Pie charts of the total occurrences in M22TBOA04 and F26NMAA26 repertoires. The highest percentage of rolled /r/s among the whole young group (66%) is the result, as admitted by speaker

M22TBOA04, of a conscious choice as a marker of 'anti-poshness'. F26NMAA26 never adapted the new sound into his speech simply because she is totally uninterested in changing the language she learnt at home.

11.3. On the theories of language change: regular vs. irregular

Possibly the earliest trend in Historical Linguistics was represented by the studies of the School of the Neogrammarians. This group of German scholars held a very strong methodological claim about sound change. According to their perspective it is, among others, blind, regular, lexically abrupt, and phonetically gradual. 'When we speak of systematic effect of sound laws we can only mean that given the same sound change within the same dialect every individual case in which the same phonetic conditions are present will be handled the same' (Paul 1880:69) translated in (Hale 2003:343). The development of the quantitative analysis of synchronic data, however, arose a different picture of the transformation of the sounds of a language, so that the Neogrammarian hypothesis was disproved.

Firstly, changes in progress have been described by Labov in *New York* and *Martha's Vineyard* (early 1960s). Wang (1969) did the same for Chinese dialects; the argument that they succeeded in describing the changes while they were taking place is just enough to prove that sound change is not blind nor lexically abrupt. That the change is not systematic was noticed also by Sommerfelt (1962) referring to a synchronic survey of Welsh dialects, which showed instances of the irregular loss of a fricative before semivowel throughout the lexicon.

The outcomes provided by this research indeed corroborate the view of *lexical diffusion*. There are no speakers with a perfect total score of approximants. Instead, even those speakers showing a heavy usage of approximants sometimes betray a conflicting sound. For instance, speaker F24MMOA06 (with one of the highest score of approximants, more than 90%) gave two realisations of [ɹ] and two of [r] in the same linguistic environment /ˈɔr/. This could mean that the change, interpersonally, is either not completed yet or completed with residual forms. Since this reasoning may be applied throughout the whole sample of speakers, we can therefore once again support the hypothesis that sound change is indeed not blind but observable and irregular because '[i]f sound change is taken to be lexically gradual, its course is neither inevitable nor inexorable, even after it has begun' (McMahon 1994:53).

The *phonetically gradual* pattern of change theorised by the Neogrammarians is, according to my opinion, a more complex issue to handle. Apparently, a sort of gradual change may be seen through the data. However, before introducing this point, I need to stress some limitations of the entire work. These mainly concern the settings for the analysis; that is, the cut-off points are just conventional, the computation of the mean formants considering only a 50% of the entire /r/ segment is also conventional, and the placement of the borders in tier 1 is done by eye so that a certain degree of error is expected. Bearing in mind all these limitations, if we look back at table 10.2

we see, for example, that speaker F51T shows a few approximants, no trills at all, and a very high number of vowels (more than one half of the entire speech material). My hypothesis, which cannot be tested here, is that she is one example of *young* old-speakers who consciously switched her speech (age-grading) towards the more prestigious and young-like variant of nowadays D.o.A.. Crucially, she fails in reproducing the 'exact' sound. The fact that she scored zero in trills and only one in fricatives out of more than 200 items clearly means that she avoids these sounds (more old-like ones). On the other hand, however, she got a very low percentage of approximants (around 16%). The large number of vowels and [ə]s, proves in my opinion, both the limitation of the work (so many vowels for a /r/ sound seem unlikely) and that she is struggling at imitating that sound that she feels as more prestigious. Yet, she is *gradually* approaching it, failing at the moment, at producing a 'proper' approximant with a 'proper' low F_3. As I said, this is just a hypothesis that should be tested in the future. If it were true, however, it would bring back part of the Neogrammarians thoughts as valuable theories for the studies of the discipline.

11.4. On the theories of language change (2): from above vs. from below

A change spreads, from person to person. The very first moment of the very first change is indeed impossible to detect. However, things have been improved and, as said many times already, a change can be detected and observed. The pessimistic view of the earliest Linguists was discarded: 'No records have ever been kept of these first beginnings of regular changes of sound… We know that English *wah* has changed to *waw*, and we can give approximate dates for some stages of this process; but we do not know when or where or in whose pronunciation the first impulse towards the change occurred' (Sturtevant 1917/1961:82).

Another insight of the change that can be gained from its properties is the level of consciousness of the speakers towards change. Needless to say, this likely affects the way the change spreads. Accordingly, two alternatives are at hand: changes may be *above* the level of consciousness so people are realising that the change is under way or *below* it, that is the change is not noticed.

In the conclusions on research on the front approximant in Dutch, van Bezooijen et al. (2002) consider the change of [r] > appx as a one from *below*. On the basis of our data we are forced to disagree with their stance: all the findings seem to point at the opposite evaluation that is a change from *above*.

Firstly, most of the participants were aware of the variability of this sound and they shared the belief that it features a typical accent of a certain network of people (the mid-high, prestigious, posh network). Therefore, the degree of consciousness seems rather developed in the speakers' minds. Decisively, they even react at it through either encouragement (most of them) or abandon

(M22TBOA04 and F26NMAA26): this indeed means that speakers noticed the change otherwise, how could they possibly have reacted to it?

Secondly, it looks that some young speakers changed their early accent when they became adult. M22TBOA04 witnessed his double change: original first acquired [r] > [ɹ] (because of his mates' partnership) > [r][ɾ] (because of his rejection to the network once he reached adulthood). Moreover F60OPHDOA34's children present patterns very different form their mother. This seems to clash with van Bezooijen et al. statement '[the change is] acquired by children from age 4'[11] (van Bezooijen, Kroezen et al. 2002:9).

Thirdly, it seems that it is not a change that occurred out of the blue, but was already in place (in the *Het Gooi* region) and that has been borrowed and slowly accepted in the D.o.A. - this is typical of a conscious change through language contact.

12. Conclusion

The main aim of the present study was to attain insight into the variation of the postvocalic /r/ in Dutch spoken in Amsterdam. The result confirmed the expectations; an ongoing phonetic change is taking place. The comparison between the old group repertoires and the young groups' showed that a broad set of sounds (trills, taps, fricatives and vowels) is apparently leaving pace to a stage of more reduced allophony in which an approximant /r/ is taking over.

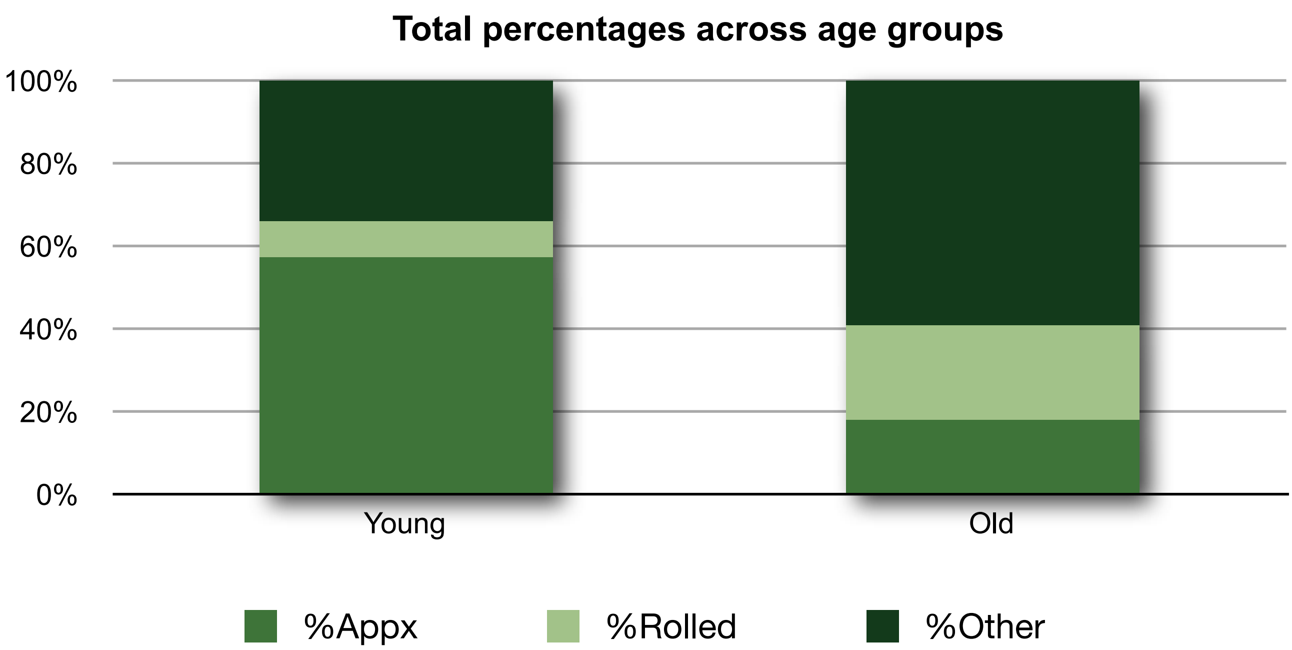

Fig. 12.1. Graph bars. The means (in percentages) of the total occurrences of approximants, rolled and other sounds across the two age groups.

11 The design of her work was extremely different from this one: the speakers were raised in two different cities of the country but not in Amsterdam and the young group was made up by children from 10 to 13 years.

Forty speakers (unevenly distributed in 13 girls, 10 boys, 9 women and 8 men) were asked to read out a word list and to speak up what they saw in an elicitation task with pictures. In total, the young group scored a 57% in the use of rhotic approximants whereas the old one only scored 18%. Conversely the average of the occurrences of the rolled /r/s among the young group is just 8% and for the old 18%. A third part is represented by all the other allophones (fricatives, vowels and zero).

The results shown above, corroborated by the statistical analysis, are consistent with the hypothesis: the sound change is undoubtedly going on. Other final remarks are: the apparent-time construct, which is the model that has been followed throughout this work, succeeded in proving that what was an impressionistic observation is really a change in Amsterdam. Apparently, women are not leading the change because the test on the mean difference between the genders of the young group gave not significant results (so they are using the new sound at the same rate). The responsible for change are the young adults that are spreading the change through a positive social evaluation. The change is from *above*, that is a change from above the level of consciousness. The initiation of the change seems to be due to contact with another dialect/area of the Netherlands. The change is indeed gradual and irregular.

The analysis of variance also showed that the production of the new sound is conditioned by the phonological context in which the postvocalic /r/ is set: stressed vowels and back vowels seem to foster the realisation of the sound as an approximant.

Along side with the change of postvocalic /r/, the [ə] epenthesis in suitable context such as /rk/, /rf/ and others, which is common in Dutch, seems to recede. The two phenomena are tied by a significant negative correlation.

Some of the more relevant limitations of the work were already mentioned in section § 10.2.5. and in section § 11.3.: they are concerned with the selection of the informants and with the conventional settings of the analysis. Throughout the accomplishment of this work more than one problem was faced. However, the greatest was represented by the lack of precise phonetic cues for discriminating sounds. This even influences our research question. The earliest hypothesis, as a matter of fact, was deeper than the one presented here. In fact, we aspired to prove that the change is not simply towards a rhotic approximant but to a *retroflex* approximant [ɻ]. However, the problem was lying on the phonetic ground: an explicit set of phonetic cues qualified to single this sound out of others, especially other approximants in general, were not available (recall section § 9.3.1.). Although we fixed the 450 Hz F_3-F_2 difference for differentiating the two sounds, we cannot utterly rely on this parameter without further evidence. As a consequence, we considered for the statistical computations the total percentages of *rhotic approximants* regardless of their point of articulation. It was due to the lack of agreed parameters that the research question was restated and the target slightly widened.

An interesting future research would therefore be based only on phonetic grounds and would seek the missing cues in order to endow future acoustic analysis with a powerful tool for an unbiased, automatic discrimination of sounds.

Abbreviations

appx	approximant sound
BV	back vowel
C	consonant
CV	central vowel
D.o.A.	Dutch of Amsterdam
FV	front vowel
IPA	International Phonetic Alphabet
M	mean
ms	milliseconds
SD	standard deviation
sec.	seconds
V	vowel
VV	long vowel
ˈV	stressed vowel
˘V	unstressed vowel
#	word boundary

Appendix - word list

Deel 1: Bloem, duur, bier, ogen, kor, heksentoer, onmiddellijk, goor, drukken, concours, halfbroer, koopje, sliert, jus, dikwerf, donderaar, werk, koningin, bewoner, draad, wekken, souvenirs, tijdsduur, sar, landheer, ter, bijna, giert, prei, mededelingenbord, mirk, parameter, appelsap, tsjirp, hier, tempo, gediplomeerd, soepjurk, trein, sluitspier, storm, goed, taart, baby, gebruiken, waar, massamoord, kaart, drossaard, kijken, idealiter, tabak, factor, vers, gaan, loopoor, sport, toer, raam, vir, gedragsleer, je kirt, kantoor, agressor, ik moet, broer, por, hond, weert, retours, spreken.

Deel 2: Ik mor, gashaard, avond, winter, schrijven, mond, meegesleurd, saldo, sfeer, pin, lor, lijkkleur, natuurlijk, boer, hij pleurt, bouwboer, slagbeurt, bus, bankier, paspoort, horen, heetgebakerd, toekomst, Dollard, augurk, erbij, meer, bioscoop, zuur, voetbal, reguliers, Kalkar, spier, kuur, punt, bangerd, melkzuur, keer, melk, hij meurt, lijnvaart, twee, normaliter, drukpers, louvredeur, pauze, buur, kans, dikoor, storen, vier, bakker, altijd, hoofdinspecteur, kwart, smaldier, derde, duurs, woord, meisje, winkel, droogschuur, zeur, kunst, maart, radar, praten.

Deel 3: Plaatsen, uit, kleur, voor, afstudeert, kip, acteur, schoonmaakbeurt, weg, laars, ophalen, medeauteur, hebben, causeur, poot, jaar, slank, per, intussen, lommerd, sterk, blauw, ver, land, soort, borst, klop, lomperd, ui, jaloers, gekir, flapoor, Geert, boterzuur, betalen, hoer, hemd, honger, minnaar, schools, letterkeer, wijn, koers, lakmoespapier, stuurs, ontbijt, solotoer, leg, dar, geld, hexameter, censor, omleiding, linnenpers, vis, berg, klein, roert, donor, echt, meneer, Osdorp, op, gestuurd, reactie, huis, hart, katje, stadskantoor, zijn.

Deel 4: Doperwt, spullen, her, geelgors, feestdag, bark, bloed, dollar, kurk, Leiden, war, mooi, arm, spurt, reservatie, gangboord, groeispurt, honing, smaakje, uitvragen, baljurk, tuinhark, televisie, ik kir, schepbord, boscultuur, aanleggen, halsboord, ei, Mir, kunsthars, ook, slurp, collega, sitar, eindspurt, taxi, lijkbaar, uw, dir, Bernard, openlijk, schaamhaar, klaar, geel, letterwoord, plek, spar, kort, nirt, rijst, les, Ruurd, morgen, boor, geweldig, ganzenveer, keurs, gewoon, koorts, aar, gevoel, schoon, dievenlantaarn, beurt, eten, verkleurd, schutdeur, gedaan.

Deel 5 (elicitation task): Roer, kilometer, ijsbeer, kerk, kaars, mier, beurs, spoor, honderd, koord, ster, dansvloer, deur, iers, dekstier, postkantoor, kar, uur, water, professor, dagkaart, ijskar, park, huiswerk, hor, dorp, kroonkurk, nieuwjaar, jurk, lampzwart, stadspoort, gefeliciteerd, schaar, buurt, Dirk, spreekbeurt, boers, beer, schietsport.

References

Aitchinson, J. (1991). Language change: progress or decay? Cambridge, Cambridge University Press.

Bailey, G. (2002). Real and Apparent Time. The Handbook of Language Variation and Change. J. K. Chambers, P. Trudgill and N. Schilling-Estes. Oxford, Blackwell Publishing.

Bailey, G., T. Wikle, et al. (1992). "The apparent time construct." Language Variation and Change **3**: 241-264.

Boberg, C. (2004). "Real and Apparent Time in Language Change: Late Adoptions of Changes in Montreal English." American Speech **79, 3**: 250-269.

Boersma, P. and D. Weenink (2011). Praat: doing phonetics by computer [Computer Program].

Booij, G. (1995). The Phonology of Dutch. Oxford, Oxford University Press.

Bussmann, H. (1996). Routledge Dictionary of Language and Linguistics. London and New York, Routledge.

Campbell, L. (2004). Historical Linguistics - An Introduction. Edinburgh, Edinburgh University Press.

Chambers, J. K. (2002). Patterns of Variation including Change. The Handbook of Language Variation and Change. J. K. Chambers, P. Trudgill and N. Schilling-Estes. Oxford, Balckwell.

Collins, B. and I. M. Mees (2003). The Phonetics of English and Dutch. Leiden - Boston, Brill.

Crystal, D. (2008). A Dictionary of Linguistics and Phonetics. Oxford, Blackwell Publishers.

Ferrand, C. T. (2002). "Harmonics-to-Noise Ratio: An Index of Vocal Aging." Journal of Voice **16**(4): 480-487.

Hale, M. (2003). Neogrammarian Sound Change. The Handobook of Historical Linguistics. D. J. Brian and R. D. Janda. US, UK, Australia, Germany, Blackwell Publishing.

Hamann, S. (2003). The Phonetics and Phonology of Retroflexes. Utrecht, LOT.

Hayward, K. (2000). Experimental Phonetics. England, Pearson Education.

Heffernan, K. (2004). "Evidence from HNR that /s/ is a social marker of gender." Toronto Working Papers in Linguistics **23**(2): 71-84.

Hockett, C. F. (1958). A Course in Modern Linguistics. New York, Macmillan.

Irvine, J. (2001). Style as distinctiveness: The culture and ideology of linguistic differentiation. Style and Sociolinguistic Variation. P. Eckert and J. Rickford. Cambridge, Cambridge University Press**:** 21-43.

Labov, W. (1994). Principles of Linguistic Change - Vol. 1 Internal Factors. Oxford, Balckwell.

Ladefoged, P. (2001). Vowels and Consonants - An Introduction to the Sounds of Languages. Malden (USA) Oxford (UK), Blackwell Publishers Ltd.

Ladefoged, P. and I. Maddieson (1996). The Sounds of the World's Languages. Oxford, Blackwell Publishers.

Lass, R. (1984). Phonology. An introduction to basic concepts. Cambridge, Cambridge University Press.

Lindau, M. (1985). The story of /r/. Phonetic Linguistics. V. A. Fromikin. Orlando, Academic Press**:** 157-168.

Maddieson, I. (1984). Patterns of Sounds. Cambridge, Cambridge University Press.

McColl Millar, R. (2007). Trask's Historical Linguistics. London, Hodder Education.

McMahon, A. (1994). Understanding Language Change. Cambridge, Cambridge University Press.

Meertens, P. J. (1938). "De radio en de uitspraak van het Nederlands." Logopaedie en Phoniatrie **10**: 53-56.

Nooteboom, S. G. (1972). Production and Perception of Vowel Duration. A Study of the Durational Properties of Vowels in Dutch. Utrecht, Utrecht. **Phd dissertation**.

Paul, H. (1880). Prinzipien der Sprachgeschichte. [2nd edn Halle: Niemeyer, 1886; 5th edn 1920. Repr. Tübingen: Niemeyer, 1970. 2nd edn trans. Herbert Augustus Strong as Principles of the History of Language. New York, Macmillan, 1890; 1st edn 1888–9. Repr. College Park, MD: McGrath, 1970.

Saussure de, F. (1915/1959). Cours de Linguistic générale. Paris, Payot, 1915.

Scobbie, J. M. and K. Sebregts (2010). Acoustic, Articulatory, and Phonological Perspectives on Allophonic Variation of /r/ in Dutch. Interfaces in Linguistics. R. Folli and C. Ulbrich. Oxford, Oxford University Press. **15:** 257.

Sebregts, K. (2011). Amsterdam.

Smakman, D. (2006). Standard Dutch in the Netherlands : a Sociolinguistic and Phonetic Description. Utrecht : LOT.

Sommerfelt, A. (1962). Diachronic and Synchornic Aspects of Language. The Hague, Mouton.

Sturtevant, E. H. (1917/1961). Linguistic change. Chicago, Chicago University Press.

van Bezooijen, R., S. Kroezen, et al. (2002). "Front approximant /r/ A new and vigorous change in Dutch." Linguistics in the Netherlands: 1-11.

van de Velde, H. (1996). Variatie en verandering in het gesproken Standaard-Nederlands (1935-1993). Nijmegen, Proefschrift Katholieke Universiteit Nijmegen.

van de Velde, H., R. van Hout, et al. (1997). "Watching Dutch change: A real time study of variation and change in standard Dutch pronunciation." Journal of Sociolinguistics **1**(3): 361-391.

van den Heuvel, H. and C. Cucchiarini (2001). /r/-deletion in Dutch: rumors or reality? 'r-atics. Sociolinguistic, phonetic and phonological characteristics of /r/. H. van de Velde and R. van Hout, Etudes & Travaux 4**:** 185-198.

Verstraeten, B. and H. Van de Velde (2001). Socio-geographical varition of /r/ in standard Dutch. 'r-atics. Sociolinguistic, phonetic and phonological characteristics of /r/. H. van de Velde and R. van Hout, Etudes & Travaux 4**:** 45-61.

Wang, W. e. (1969). "Competing changes as a cause of residue." Language 45: 9 - 25.

Weireich, U., W. Labov, et al. (1968). Empirical Foundations for a Theory of Language Change. Directions for Historical Linguistics. W. P. Lehmann and M. Yakov. Austin, University of Texas Press**:** 95-195.