/r/ as language marker in bilingual speech production and perception

Constantijn Kaland, Vincenzo Galatà, Lorenzo Spreafico, Alessandro Vietti

ALPs - Alpine Laboratory of Phonetics and Phonology, Free University of Bozen – Bolzano, Italy

{constantijn.kaland, vincenzo.galata, lorenzo.spreafico, alessandro.vietti} @unibz.it

Abstract

Across languages of the world /r/ is known for its variability. Recent literature incorporates sociolinguistic factors, such as bilingualism, in order to explain /r/ variation. The current study investigates to what extent /r/ is a marker of a bilingual's dominant language. Specifically, the effects of several sociolinguistic and phonotactic factors on the production and perception of /r/ are investigated, such as the bilingual speaker's linguistic background, the language spoken as well as syllable position and place of articulation. To this end a reading task is carried out with bilingual speakers from South Tyrol (Italy). The major languages spoken in this region are Tyrolean (German dialect) and Italian. The recorded reading data is subsequently used in a perception experiment to investigate whether South Tyrolean listeners can identify the dominant language of the speaker on the basis of the presence of /r/ and the /r/ variant. Results show that listeners can identify the dominant language of the bilingual speakers on the basis of /r/. Specifically, the more Italian dominant the sociolinguistic background of the speaker, the more /r/ is produced frontally and the more that speaker is perceived as Italian dominant.

Index Terms: speech production, speech perception, variation, bilingualism, articulation

1. Introduction

The variation of /r/ has been widely studied. Research has focused on phonetic and phonological (i.e. segmental) properties of /r/ in order to structure the rich amount of variants ([1],[2],[3],[4]). Several other studies explain variation of /r/ by sociolinguistic factors, such as bilingualism ([5],[6],[7]). The literature generally lacks perceptual evidence on the effects of sociolinguistic factors on /r/ variation [8]. A discussion of the most relevant studies is given in this section.

It has been a difficult task to find a single phonetic property of /r/ that unifies the variants into a phonological class of rhotics. Acoustically, formant changes in F2 and F3 have been suggested as a cue for rhoticity [9], but it also has been shown that rhoticity is perceived when F3 is absent [10]. Because of the lack of a unifying feature, phonological models tend to include more features and describe rhotics in terms of parametric relations ([1],[3]). These models, however, still rely to a high extent on phonetic features of /r/.

The recent literature incorporates sociolinguistic variables to study /r/ variation [11]. This approach is applied to, among other languages, Persian [12], Arabic-English [6], Canadian French [13] and Dutch [4]. For example, in Standard Dutch it is common to find variation in the use of /r/ due to changes in speaking style [5], regional language variety or sociolect [4].

The study of bilinguals appears to be particularly useful to investigate /r/ variation. That is, the variation in bilinguals is often unique due to the specific combination of languages mastered by the speaker. In addition, bilingualism is often the result of a particular sociolinguistic background of the speaker. For example, Arabic-English fluent bilinguals are able to produce the /r/ variants that belong to either English or Arabic [6]. Nevertheless, code-switching and borrowing of variants occur in specific settings in order to retain communicative efficiency. This result is explained as a sociophonetic competence of the speaker, rather than a natural consequence of language interference in bilinguals [6].

Recent work concentrates on /r/ variation among bilinguals in the Italian region of South-Tyrol ([7],[11]). The major languages spoken in this region are Tyrolean, a Bavarian dialect of German [14], and Italian. Most South Tyroleans are proficient bilinguals, as Italian and German are used in public life, politics and education. South Tyrol is a particularly interesting region because of the political status of the respective languages. In order to divide jobs in the public sector, the regional government asks South Tyroleans to declare their loyalty to one of the speaking communities. This declaration is argued to segregate the language groups ([15],[16]). Among the South Tyrolean /r/ variants there are trills, flaps/taps, fricatives and approximants. Alveolars are most often produced by Italian dominant speakers, whereas uvulars and vocalizations are most often produced by Tyrolean dominant speakers [7]. This is in line with standard /r/ productions in Italian [17] and German [18].

While the acoustic and articulatory variation of /r/ is studied widely, less is known about the relevance of /r/ in perception. It is shown that listeners are able to identify a language mainly on the basis of segments and to a lesser extent on the basis of prosody [19]. Furthermore, it is known that phoneme boundaries vary depending on the language a bilingual is listening to [20] and that fine grained acoustic cues such as formant changes in /r/ are used by listeners to identify regional dialects [21]. Given the sociolinguistic significance of /r/, particularly in South Tyrol, this study hypothesizes that /r/ is a marker of the dominant language of a bilingual speaker from this area. Specifically, it is expected that listeners are better able to identify the speaker's dominant language when /r/ is present compared to when /r/ is absent. Furthermore, it is expected that back /r/ is associated with Tyrolean and that front /r/ is associated with Italian. To test these hypotheses, a reading task is carried out among bilingual speakers from South Tyrol. The degree of bilingualism is assessed in a questionnaire prior to the reading task. The recorded words in the reading task are then used as stimuli in a perception experiment to investigate to what extent listeners use /r/ as a cue to identify the dominant language of the speaker.

2. Methodology

2.1. Questionnaire

In the questionnaire participants ($N = 20$) judge their own proficiency (scale 0-10) for both Tyrolean and Italian and fill out ten questions for which they can report whether they speak Tyrolean or Italian, or both. The ten questions elicit which language participants speak/spoke during several stages of their life (5 Qs), which language(s) they speak with family members or friends (3 Qs) and which language(s) they speak at home or at work (2 Qs). The degree of bilingualism is calculated for each participant on the basis of a difference score, which is obtained by adding up the self-declared grade and the amount of times the participant mentioned the respective language as an answer to the linguistic background questions. In this way, the highest possible score per language is 20. The difference score is calculated by subtracting the Tyrolean score from the Italian score. Negative difference scores indicate the degree to which the participant's dominant language is Tyrolean and positive difference scores indicate the degree to which the participant's dominant language is Italian. Balanced bilinguals are expected to have a difference score around zero. Henceforth we refer to the scalar difference score as *degree of bilingualism*. The categorical distinction between Tyrolean dominant or Italian dominant is referred to as *speaker's dominant language* and depends on whether the difference score falls within the range -20 to 0, or 0 to 20 respectively.

2.2. Reading task

Each participant reads two lists which consist of 42 carrier sentences each; one list in Italian and one list in South Tyrolean (*target language*). The carrier sentence is either the phrase "ho detto [word]" (Italian) or "i hon gsog [word]" (Tyrolean), meaning "I have said [word]". In this way the recorded word has a constant sentence final position. All words have one or two syllables and are common words in Italian or South Tyrolean. The initial and final word of each list is a dummy word to prevent effects of a list intonation. From the remaining 40 words, half are target words (containing /r/) and half are control words (not containing /r/). The /r/ occurs once in each target word. Each target word pairs minimally with a control word such that the /r/ in the target word is either substituted or deleted in the control word. The substitution or deletion occurs either in syllable onset or syllable coda (Table 1) in such a way that all combinations are distributed evenly in the word list.

Table 1. Examples of four types of minimal word pairs in each target language.

Target language	Substitution		Deletion	
	Onset	Coda	Onset	Coda
Tyrolean	*brechn*	*forzn*	*schraibn*	*oaner*
	blechn	*fotzn*	*schaibn*	*oane*
Italian	*rima*	*barca*	*frusta*	*corsa*
	cima	*banca*	*fusta*	*cosa*

2.3. Reading task procedure

Before the start of the experiment, participants are given instructions about the course of the tasks. They familiarize themselves with the target words before the start of the reading task in order to prevent reading errors. Participants carry out the reading task in a laboratory with no background noise. They read the carrier sentences one by one from a computer screen. The carrier sentences appear in a timed manner on the screen to keep the reading pace constant. The time between each sentence is 4.5 seconds. To counterbalance any effect of reading order, half of the participants start reading the Italian list and the other half of the participants start reading the Tyrolean list. During the task participants wear a headset microphone (Labtec Axis 502) connected to a laptop computer (Conexant sound card) on which the recordings are saved as mono 44100 Hz wave files.

2.4. Data labelling

All target words (containing /r/) are annotated on the basis of place of articulation as either front or back ($N = 799$). One target word could not be labelled due to a recording error. All alveolar trills, taps, fricatives and approximants as well as retroflex taps/flaps are labeled "front", and all uvular trills, taps, fricatives and approximants as well as all vocalized /r/ sounds (all are open back vowels) are labeled as "back". A trained phonetician labelled the target words on the basis of audio and visual (spectrogram) inspection using Praat [22]. A representative subset of 100 words was then double checked by two other trained phoneticians. Among the three raters there was an almost perfect agreement (Fleiss Kappa = 0.85, [23]). For statistical analyses proportions of back /r/ are taken, such that 1 corresponds to the occurrence of back /r/ and 0 corresponds to the occurrence of front /r/ (see section 2.8, Table 2).

2.5. Perception experiment

Participants ($N = 25$) fill out a questionnaire identical to the one described in section 2.1. From all the words recorded in the reading task ($N = 1680$) a subset is selected for use in the perception experiment. Dummy words and words containing errors (reading or recording errors) are left out. The subset consists of a balanced set of the four types of minimal word pairs (Table 1). For each speaker in the data collection ($N = 20$) all four types of word pairs are selected per language, resulting in a total of 320 words in the perception experiment.

Participants are instructed to indicate for each stimulus word how Tyrolean or how Italian the word is pronounced. There are two caveats that make the perception experiment prone to misunderstanding by the participants. First, participants know the language of each word (this is indicated) and could therefore base their judgment on the language of the word. For this reason, participants are instructed to base their answer on the pronunciation of the words, not on the language of the word. Second, participants could memorize and base their judgment on the voice of the speaker. Therefore, participants are told that all speakers are actors who use an Italian or Tyrolean accent, which could differ across stimuli. By reevaluating each stimulus word, participants are more likely to focus on the pronunciation and thus take into account the segmental makeup of the word (and thus /r/) in each judgment they give. In addition, an explicit instruction is given not to base the judgment on the voice of the speaker.

2.6. Perception experiment procedure

The perception experiment is designed using WWstim [24], a cgi-script to carry out linguistic experiments via online html pages. Stimulus words ($N = 320$) are presented in a random order to avoid that the words are perceived as pairs. For each participant a different random order is used to minimize learning effects and minimize effects of participant fatigue. Participants see each stimulus word written on their screen. The language of the word is also indicated in order to avoid misinterpretation of words that might occur in both languages (i.e. *bar*). For each stimulus, participants click on a play-button to listen to the word. Thereafter, participants are asked to indicate how Tyrolean or how Italian the word is pronounced. This indication is given using a slider, i.e. an interactive element in the html-interface of the experiment that can be moved to the left or the right within a given range. The position of the slider corresponds to a value on a scale of 0 to 100, where a response value of 0 corresponds to the most Tyrolean perception and a response value of 100 corresponds to the most Italian perception. The default position of the slider is the midpoint of the scale (response value 50). No feedback is given about the participants' performance during the experiment. Participants carry out the experiment behind a computer or laptop using earphones. The duration of the perception experiment is approximately 30 minutes, including a short break after completion of half of the stimuli. Response values are collected on a web-server. A match score is calculated to investigate whether participants indicated the speaker's dominant language correctly (0-50 for Tyrolean, 50-100 for Italian) or not.

2.7. Participants

In total 20 participants completed the reading task (9 males, 11 females, M_{age} = 27.1 years, age range: 19-42) and 25 participants completed the perception experiment (7 males, 18 females, M_{age} = 26.8 years, age range: 20-39). All participants were recruited as bilingual speakers living in South Tyrol without problems in speaking, reading or hearing. After participating they were all debriefed about the actual purpose of the study. Participants in the reading task and the participants in the perception experiment were recruited on the basis of similar degrees of bilingualism (M_{read} = 2.5, SD_{read} = 10.1; M_{perc} = -2, SD_{perc} = 9.5).

2.8. Statistical analyses

Statistical tests are carried out separately on the proportions of back /r/ (production) and response values (perception). Thereafter, a correlation analysis is given on both measures.

A univariate analysis of variance is carried out with the proportion of back /r/ in the target words as dependent variable and speaker's dominant language (2 levels: Tyrolean, Italian), target language (2 levels: Tyrolean, Italian) and syllable position (2 levels: onset, coda) as independent variables.

Two independent sample t-tests are carried out on the match scores with the presence of /r/ in the stimulus word (2 levels: present, absent). One test is carried out on the data of the Tyrolean dominant speakers and one test is carried out on the data of the Italian dominant speakers.

Furthermore, Pearson correlation measures are calculated between the degree of bilingualism of the speaker, the mean proportion of back /r/ (production) and the mean response value given to that speaker (perception).

3. Results

3.1. Production

There is a significant effect of speaker's dominant language [$F(1,791)$ =184.55, $p < .001$, $\eta_p^2 = .19$] in that Tyrolean dominant speakers produce more back variants ($M = .61$) than Italian dominant speakers ($M = .26$). Target language is significant [$F(1,791) = 440.50$, $p < .001$, $\eta_p^2 = .36$] in that speakers produced more back variants for Tyrolean words ($M = .71$) compared to Italian words ($M = .16$). A significant effect of syllable position [$F(1,791) = 13.94$, $p < .001$, $\eta_p^2 = .02$] reveals that there are less back variants produced in onset position ($M = .39$) than in coda position ($M = .48$). The interaction between speaker's dominant language and syllable position [$F(1,791) = 11.21$, $p < .01$, $\eta_p^2 = .01$] reveals that the difference between produced back variants in onset versus coda position is larger in Italian ($M_o = .17$, $M_c = .35$) than in Tyrolean ($M_o = .60$, $M_c = .62$). The interaction between target language and syllable position [$F(1,791)$ =9.95, $p < .01$, $\eta_p^2 = .01$] reveals that the difference between produced back variants in onset versus coda position is larger in Tyrolean words ($M_o = .62$, $M_c = .80$) than in Italian words ($M_o = .16$, $M_c = .17$). The three-way interaction between speaker's dominant language, target language and syllable position [$F(1,791) = 12.54$, $p < .001$, $\eta_p^2 = .02$] shows that generally more back variants are produced in coda position when compared to onset position. This difference is the largest when speakers use their non-dominant language, specifically when Italian dominant speakers produce Tyrolean words (Table 2).

Table 2. Mean proportions of back /r/ (*SD*) in the target words as a function of speaker's dominant language, target language and syllable position.

Speaker's dominant language	Tyrolean		Italian	
	Onset	Coda	Onset	Coda
Tyrolean	.90 (*.30*)	.90 (*.30*)	.31 (*.47*)	.33 (*.47*)
Italian	.33 (*.47*)	.69 (*.47*)	.00 (*.00*)	.01 (*.10*)

3.2. Perception

Table 3. Mean match scores (*SD*) and t-test results per speaker's dominant language and /r/ presence.

Speaker's dominant language	/r/ present	/r/ absent	$t(3998)$	*p*-value
Tyrolean	.70 (*.46*)	.64 (*.48*)	4.15	< .001
Italian	.71 (*.45*)	.67 (*.47*)	2.94	< .01

The independent samples t-tests reveal that participants in the perception experiment are significantly better able to indicate the dominant language of the speaker for words that contains /r/ compared to words that do not contain /r/. This effect is found for both speaker groups (Table 3).

Table 4 shows that the highest response values are obtained for Italian dominant speakers producing an Italian word with front /r/ and the lowest values are obtained for Tyrolean dominant speakers producing a Tyrolean word with back /r/. Values close to 50 show that participants indicate Tyrolean and Italian almost equally as dominant language.

This is particularly the case for Tyrolean dominant speakers producing a Tyrolean word with front /r/ ($M = 48.13$).

Table 4. Mean response values (*SD*) of target words only as a function of speaker's dominant language, target language and place of articulation.

Speaker's dominant language	Tyrolean		Italian	
	Front	Back	Front	Back
Tyrolean	48.13 (*22.52*)	10.93 (*8.43*)	69.91 (*16.89*)	13.64 (*12.30*)
Italian	76.72 (*12.35*)	34.87 (*11.94*)	82.03 (*10.52*)	18.40 (*17.23*)

3.3. Production-perception

Figure 1 shows high correlations between degree of bilingualism and production [$r(18) = -.72$, $p < .001$]; between degree of bilingualism and perception [$r(18) = .89$, $p < .001$]; and between production and perception [$r(18) = -.92$, $p < .001$].
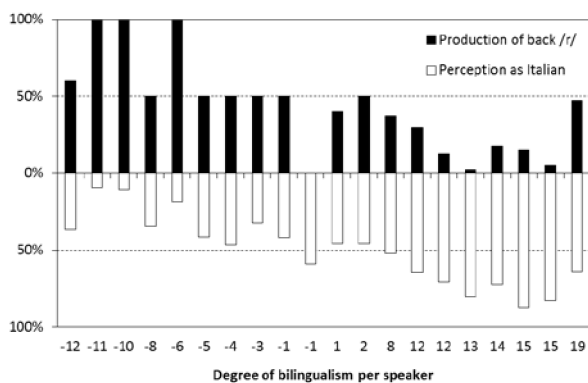

Figure 1. Production and perception measures against degree of bilingualism per speaker ($N = 20$).

4. Conclusion and discussion

The results of the analyses on the production data show that Tyrolean dominant speakers mostly produce /r/ variants with a back place of articulation, whereas Italian dominant speakers mostly produce /r/ variants with a frontal place of articulation. Back variants are more often produced in coda position than in onset position. In the current results the interaction between speaker's dominant language and syllable position shows that the difference in the use of back variants between onset and coda position is larger for Italian dominant speakers than for Tyrolean dominant speakers. This result can be explained by the tendency of Italian dominant speakers to vocalize /r/ in coda position [7]. Vocalization, then, is a relatively easy way for Italian dominant speakers to "sound" Tyrolean when they are not able to produce back /r/. This conclusion is supported by post-hoc inspection of the production data (cf. Table 2), which reveals that back /r/ produced by Italian dominant speakers generally occurs in Tyrolean words.

The results of the perception test show that listeners are better able to indicate the dominant language of the speaker when hearing /r/. This result shows that the /r/ segment is indeed a marker of the speaker's dominant language. Participants in the perception experiment were able to correctly identify the dominant language of the speaker also when /r/ was absent. This reveals that besides /r/ other acoustic cues affected participants' responses. Analysis of the target words (containing /r/) reveals that listeners use the place of articulation of /r/ as a cue to identify the speaker as either Tyrolean (back) or Italian (front). The strength of this cue can be seen when considering the second lowest ($M = 13.64$) and second highest ($M = 76.72$) response values. These values are found for Tyrolean speakers producing back /r/ in an Italian word and for Italian speakers producing front /r/ in Tyrolean words. This shows that a deviation from the expected production (Tyrolean: back, Italian: front) acts itself as a cue for the speaker's dominant language. The cases in which speakers produce a word in their dominant language with an unexpected place of articulation show an asymmetry. In these cases, Tyrolean dominant speakers are identified almost equally as either Tyrolean dominant or Italian dominant ($M = 48.13$), whereas Italian dominant speakers are mainly identified as Tyrolean ($M = 18.40$). This asymmetry indicates that the association of back /r/ with Tyrolean is stronger than the association of front /r/ with Italian, which is supported by the production data in that the production of back /r/ in Tyrolean dominant speakers is higher than the production of front /r/ in Italian dominant speakers.

The combined production-perception analysis indicates that place of articulation of /r/ strongly correlates [25] with the perceived dominant language of the speaker. It also shows that the assessment of degree of bilingualism in the questionnaire is a useful scale to investigate the distribution of /r/ variation (Figure 1). In both the production and perception data Tyrolean seems to be the marked language in that the presence and place of articulation of /r/ has larger effects compared to Italian. This conclusion is supported by the fact that Tyrolean has the status of a dialect, whereas Italian is recognized as an official language in South Tyrol. It is therefore more common to find a Tyrolean dominant speaker who is highly proficient in Italian than to find an Italian dominant speaker who is highly proficient in Tyrolean. Further support for this conclusion can be found in speakers who are able to produce Italian words with front /r/ and Tyrolean words with back /r/, i.e. participants who are highly proficient in both languages. This type of speaker is mainly found among the Tyrolean dominant speakers (black 50% bars in Figure 1).

In sum, the current study shows that the presence and place of articulation of /r/ are markers of the bilingual speaker's dominant language in South Tyrol. Furthermore, the degree of bilingualism, as determined mainly by the languages used in the social network of the speaker, provides an explanation for the variation in /r/. Specifically, /r/ variation in production and perception strongly relate to the degree to which bilingual speakers use their respective languages. The current study is a further step towards integrating sociolinguistic factors to model /r/ variation and future research should include more phonetic detail. It is important to note that in this study the labelling of variants was reduced to two places of articulation: front and back. This method ignores more fine-grained articulatory differences between /r/ variants, on which current phonological models rely.

6. References

[1] Lindau, M. "The story of /r/", in V.A. Fromkin [ed.], Phonetic Linguistics: Essays in honour of Peter Ladefoged, 157-168, Orlando: Academic Press, 1985.

[2] Ladefoged, P. & Maddieson, I. "The Sounds of the World's Languages", Oxford: Blackwell, 1996.

[3] Magnuson, T.J. "The story of /r/ in two vocal tracts", Proceeding of ICPhS 16, 1193-1196, Saarbrücken, 2007.

[4] Sebregts, K. "The Sociophonetics and Phonology of Dutch r", PhD thesis, LOT: Utrecht, 2014.

[5] Van de Velde, H. & Van Hout, R. "The pronunciation of (r)", in R. van Bezooijen & R. Kager [eds.], Linguistics in the Netherlands, 177-188, Amsterdam: John Benjamins, 1999.

[6] Khattab, G. "/r/ production in English and Arabic bilingual and monolingual speakers", Leeds Working Papers in Linguistics & Phonetics, 9(1): 91–129, 2002.

[7] Vietti, A., Spreafico, L. and Galatà, V. "A UTI Study on the Phonetic Allophony of Tyrolean /R/", Proceedings of ICPhS 18, Edinburgh, 2015.

[8] Drager, K. "Sociophonetic variation in speech perception", Language and Linguistics Compass, 4:473-480, 2010.

[9] Engstrand, O., Frid J. & Lindblom, B. "A perceptual bridge between coronal and dorsal /r/", in M.-J. Solé, P. S. Beddor & M. Ohala [eds.], Experimental Approaches to Phonology, Oxford: Oxford University Press, 2007.

[10] Heselwood, B. & Plug. L. "The role of F2 and F3 in the perception of rhoticity: evidence from listening experiments", Proceedings of ICPhS 17, 867-870, Hong Kong, 2011.

[11] Spreafico, L. & Vietti, A. "Rhotics: New Data and Perspectives", BU Press: Bolzano, 2013.

[12] Rafat, Y. "A sociophonetic investigation of rhotics in Persian", Iranian Studies, 43(5): 667-682, 2010.

[13] Sankoff, G. & Blondeau, H. "Instability of the [r] ~ [ʀ] alternation in Montreal French: An exploration of stylistic conditioning in a sound change in progress", in [11], 2013.

[14] Wiesinger, P., "The Central and Southern Bavarian Dialects in Bavaria and Austria", in C.V.J. Russ [ed.], The Dialects of modern German: a linguistic survey, London: Routledge, 1990.

[15] Eichinger, L., "South Tyrol: German and Italian in a Changing World", Journal of Multilingual and Multicultural Development, 23(1-2): 137-149, 2002.

[16] Meluzzi, C. "Two groups, two worlds: Italian and German in Bozen, South Tyrol", in S. Bagga-Gupta (ed.), Marginalization Processes: Studies of membership and participation across disciplines and sites, Cambridge: CSP, 2015.

[17] Canepari, L. "Dizionario di Pronuncia Italiana" (DiPI), Bologna: Zanichelli, 2009.

[18] Krech, E.M., Stock, E., Hirschfeld, U. & Anders, L-C., "Deutsches Aussprachewörterbuch", Berlin: De Gruyter, 2009.

[19] Bezooijen, R. van & Gooskens, C., "Identification of language varieties: The contribution of different linguistic levels", Journal of Language and Social Psychology, 18: 31–48, 1999.

[20] Elman, J., Diehl, R. & Buchwald, S., "Perceptual switching in bilinguals", Journal of the Acoustical Society of America 62: 971-974, 1977.

[21] Clopper, C. G., & Pisoni, D. B., "Some acoustic cues for the perceptual categorization of American English regional dialects", Journal of Phonetics, 32:111–140, 2004.

[22] Boersma, P. & Weenink, D. "Praat: doing phonetics by computer" [Computer program], v. 5.4.18, retrieved 21 September 2015 from http://www.praat.org/.

[23] Landis, J. R., & Koch, G. G. "The measurement of observer agreement for categorical data", Biometrics 33: 159–174, 1977.

[24] Veenker, T. J. G. "WWStim: A CGI script for presenting web-based questionnaires and experiments", v. 1.4.4, Utrecht University, 2003.

[25] Taylor R. "Interpretation of the correlation coefficient: a basic review", J. of Diagnostic Medical Sonography, 1:35–39, 1990.