

HOW CREDIBLE IS A NON-NATIVE SPEAKER? PROSODY AND SURROUNDINGS

Anna De Meo

University of Naples “L’Orientale”
ademeo@unior.it

ABSTRACT

In order to evaluate the segmental and suprasegmental acoustic credibility correlates in the perception of native and non-native speech, a study was conducted on native and non-native speakers (NSs and NNS) of Italian. NNSs were chosen on the basis of L2 level of competence (A2, B1, C1 of CEFR), mother tongues (Chinese, Arabic, Japanese and Vietnamese) and degree of foreign accent (mild vs. strong). Four perceptive tests were carried out, two of them based on natural speech and the other two on modified speech. The corpus, formed by bizarre-but-true news read in Italian by all the involved subjects, was organized in form of radio news magazines and was used to conduct the tests. Native Italian listeners were asked to assess the *truthfulness* of each news item and its *comprehensibility*. Results revealed that the correlation between foreign accent and credibility is delivered by comprehensibility: poor comprehensibility generally lowers the credibility level of an utterance. When comprehensibility is high, a reduced tonal range and longer silent pauses, i.e. the suprasegmental features of the utterance, determine a significant increase of trustworthiness.

Keywords: foreign accent, credibility, L2 Italian, prosody, comprehensibility.

1. INTRODUCTION

Direct personal experience and other people’s oral/written communication constitute a major source for gathering new information, although most of this content is filtered out and only what is thought to be useful is retained. One of the primary criteria used to select the collected information is its credibility. Sources are defined credible or believable when they are described as trustworthy or as having experience [10, 16, 21]. If the information is rejected, because it is judged as not credible, it is not going to be learned nor stored in the memory and it will not have any impact on one’s life. Given that credibility has such an influence on the message, it is important to understand how people decide what to believe in.

Message credibility is generally the result of interactions between source characteristics, e.g. expertise, trustworthiness, dynamism, authority, and message features as perceived by the recipient of the message. Factors that impact credibility are also strongly related to receiver’s characteristics and include cultural background, previous beliefs, age, gender, country of origin, education, etc.

Speaker’s accent, defined as the perceived segmental and prosodic deviations from “standard norms” in native and non-native speech production (regional accent vs foreign accent) [3, 11, 14, 18],

may impact his/her credibility because it can activate listeners’ positive or negative prejudices or stereotypes, i.e. preconceived beliefs about a group of people or cultural practices [8, 17].

As Italy confronts increasing diversity, primarily due to migration (150 languages spoken by more than 3.5 millions immigrants) and globalization trends, listener’s perceptions of the varieties of spoken Italian has deeply modified. Prejudice, once regarding regional varieties and dialects [1, 7, 20] has shifted towards foreign-accented varieties [4]. Foreign accented speech may positively or negatively affect social behaviours. In service contexts where visual cues are absent (call centers), customer’s perception of service quality may be unfavorably influenced by accent-related stereotype activation. Listeners’ prejudicial reactions may thus impact employments aspects [19].

To this regard, according to Lev-Ari and Keysar [9], foreign accent makes non-native speakers less truthful to listeners, not just because of prejudice towards foreigners, but rather because of the voice signal characteristics, that would cause processing difficulties. his assumption, based on the socio-psychological concept of “processing fluency”, i.e. a feeling of ease associated with a cognitive operation, which may affect the way stimuli are judged [12, 15], takes for granted that the acoustic-perceptual traits of non-native speech

signal are acoustically harder to process. The authors investigated what happens to credibility when stereotypes about non-native speakers are neutralized, by presenting foreigners as simple locutors of messages created by native speakers. They demonstrate that heavy accented speech is harder to believe, basing their observation on a corpus of trivia statements read by various non-native speakers of American-English. However, the study did not pay close attention both to the acoustic features characterising the foreign accent, and to the influence of the various L1s and speakers' L2 level of competence on listeners' perception.

Since the topic is socially very relevant, a study was designed to evaluate and assess the acoustic credibility correlates - segmental and/or suprasegmental - in the perception of native and non-native speech. In this paper, the results of four experiments performed on natural and modified read speech of native and non-native speakers of Italian are reported [5, 6, 13].

2. METHODS AND MATERIALS

2.1. Speakers and listeners

This study involved three Italian native speakers (NS) and six non-native speakers (NNS) of Italian, all university students living in the Campania Region, Southern Italy.

NNS were chosen on the basis of their level of L2 competence in Italian, their mother tongue and their degree of foreign accent (mild and strong). In order to determine the degree of the perceived accent, a "global foreign accentness rating test" was used and was administered to 70 native Italian listeners (all Campanian), who rated the degree of foreign accent of a short read text on a three-point scale (1= native speaker; 3= strong foreign accent).

The involved NNSs, all learners of L2 Italian, were:

- two advanced learners (C1 level of the CEFR), male and female, both Chinese, differing only for the degree of foreign accent (mild vs. strong);
- two intermediate learners (B1 level), both female, Chinese and Vietnamese L1s, strong foreign accent;
- two beginners (A2 level), both female, having Arabic and Japanese as L1s, strong foreign accent.

2.2. Corpus

As previously stated, the text of the message can influence credibility judgment. For this reason the corpus was formed by 12 bizarre-but-true news

from around the world, similar to the following piece of text: "A Port St. Lucie, Florida, woman ended a night of heavy drinking by stabbing a male friend four times with a long, pointy mollusk shell. The bivalve violence was apparently so brutal that cops charged Patricia Wehr with aggravated battery with a deadly weapon".

All the pieces of news were read in Italian by native and non-native speakers and were organized in radio news magazines, combining different voices in random order. All audio recordings were made in the University WebRadio recording studios.

The radio news magazines were used to carry out four perceptive tests administered to native Italian listeners, 951 university students, male and female, all Campanian in order to avoid any diatopic influence in the perceptive judgment.

Tests were proposed as a survey on media reliability, in order to avoid catching the listeners' attention on the foreign voices.

2.3. Perceptive tests

The four perceptive tests were organized into two groups: the first two tests were based on natural speech; the other two on artificially modified speech.

- **Test 1:** 4 voices (2 NSs and 2 Chinese NNSs C1 level; male and female in both groups); 301 Italian listeners.
- **Test 2:** 5 female voices (1 NS and 4 NNSs having a strong foreign accent, A2-B1 levels); 265 Italian listeners.
- **Test 3:** test 1 voices, with tonal range and silent pauses artificially increased and decreased by WinPitch [23]; 120 Italian listeners.
- **Test 4:** test 2 voices, with errors and disfluencies removed, segmental duration and tonal range cloned from the native voice using WaveSurfer 1.8.8 [22] and Praat 5.3.30 [2]; 265 Italian listeners.

Listeners were asked to assess the truthfulness of each piece of news and evaluate its comprehensibility, i.e. listener's estimation of difficulty in understanding the utterance. For the last two tests, based on manipulated speech, the degree of the perceived foreign accent was also evaluated, in order to point out any occurred changes in the listeners' judgments.

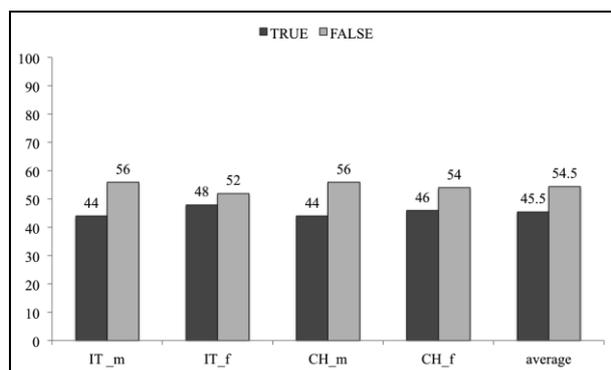
The suprasegmental features of the whole corpus were analysed using WaveSurfer 1.8.8. Speech chains duration, number of syllables per speech chain, silent pauses duration, disfluencies duration (vocalization, nasalization, interruptions, etc.), maximum and minimum f_0 per speech chain

were measured. Articulation rate (syl/s), speech rate (syl/s), fluency (syl/number of speech chains), mean duration of silences (s), percentage of silence, tonal range (in semitones) were calculated as well.

3. RESULTS

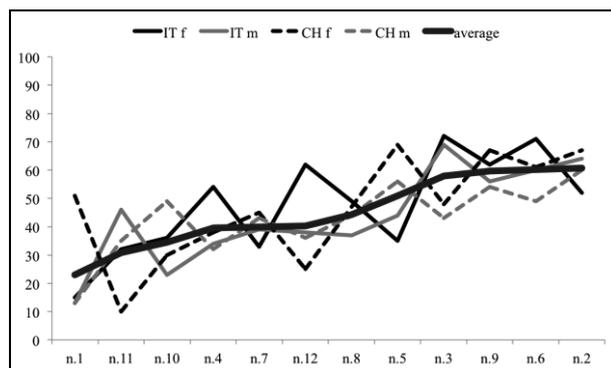
3.1. Test 1

Figure 1: Average percentage of credibility per speaker.



As expected, for this first test, due to the advanced level of competence of the two NNSs, utterances showed no disfluencies. The degree of comprehensibility was very high for both native and non-native speakers. The average credibility level was around 50%, therefore in a range of randomness. Since the two Chinese had two different degrees of foreign accent, mild and strong, it is evident that the native listeners are not influenced by this element (Fig. 1).

Figure 2: Percentage values of speakers' credibility per each piece of news.



Each piece of news seems to have its own degree of credibility, which depends on the content of the message. Some news items were rated very low on credibility (e.g. news 1 - 23% true) while others were judged as more credible (e.g. news 2 - 61% true). Nevertheless, results show that ratings are significantly different (up to $\pm 37\%$), depending on the speaker but regardless of his/her language nativeness (Fig. 2). NSs and NNSs achieved both high and low scores on the test.

Since the communication situation (i.e. formally, hyperarticulated, read speech) and the speakers' phonetic repertory remained unchanged, the impact of the segmental level on the credibility was excluded.

As for the rhythmic-prosodic features NSs showed expected average articulation and speech rate values higher than those of the NNSs. These data are basically stable for each speaker, therefore they do not justify the differences in credibility judgment (Table 1).

Table 1: Articulation rate (AR) and Speech rate (SR).

	average AR	average SR
Italian speakers	6.7	6
Chinese speakers	5.3	4.7

On the contrary, tonal range appeared to be significant. Different news delivered by the same speaker showed a credibility increase when the speaker restricted the tonal range, avoiding marked tonal peaks (Table 2). The relation between credibility and low tonal range was confirmed by data concerning different speakers uttering the same piece of news, either with different or similar rhythmic-prosodic features (Table 3).

Silent pauses seemed to play a minor role. Gender and accent, both strong and mild, played no role at all.

Table 2: Tonal range (st) and credibility values (%) – same speaker/different news.

CH_f		
	News 1	News 11
tonal range	8.4	18.8
credibility	51	10
IT_m		
	News 11	News 10
tonal range	9.65	18.16
credibility	46	23

Table 3: Tonal range (st) and credibility values (%) – same news/different speakers.

NEWS 5		
	CH_f	IT_f
tonal range	9.8 (st)	12.2 (st)
credibility	69	35
NEWS 6		
	CH_f	IT_m
tonal range	7.5	7.6
credibility	61	60

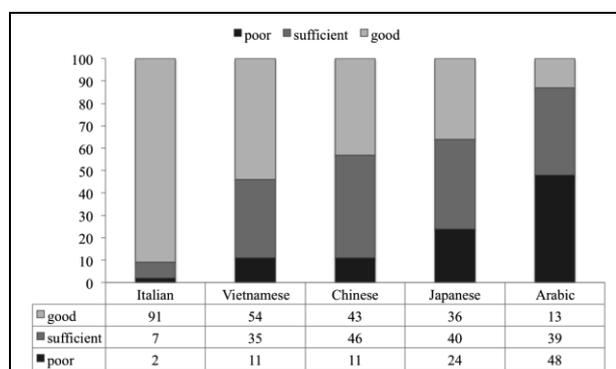
3.2. Test 2

Since gender played no role in the credibility judgment, only female voices were used for the second test. In addition, aiming at the evaluation of

possible negative influence of all the features characterizing the speech of beginners and intermediate speakers of an L2, i.e. disfluencies, wrong pauses, interruptions, anomalous tonal variation, on comprehensibility, and consequently on credibility, four A2 and B1 NNSs with strong foreign accent (L1s Arabic, Chinese, Japanese and Vietnamese) were involved for the second test. To this group a female NS was added.

Perceptive test results showed that only the NS received was considered very comprehensible (“good” 91%). Intermediate values were attested for the two B1 NNSs (“good”: Vietnamese 54% and Chinese 43%), while the A2 NNSs were judged as “poor” comprehensible (“good”: Japanese 36% and Arabic 13%) (Fig. 3). Data indicate that the degree of comprehensibility is strictly related to the level of L2 competence.

Figure 3: Comprehensibility average values (%).



When there are no comprehensibility problems the assessment of true/false is stable around 50%, as already observed in the previous test (Table 6).

Table 3: Credibility of news having “good” comprehensibility (%).

	TRUE	FALSE
Arabic	46	54
Chinese	51	49
Italian	54	46
Japanese	61	39
Vietnamese	45	55

On the contrary, when the level of comprehensibility lowers, due to anomalous acoustic traits (disfluencies, errors, etc.), the judgment of “false” rapidly increases, reaching 90% when the statement proves to be poorly understandable for at least the 40% of listeners.

Therefore, there seems to be a threshold of comprehension tolerance, i.e. a level of difficulty in understanding an utterance at which the listener’s effort to understand the message leads

him/her to believe that what he/she has just heard is not credible (Table 7).

Table 4: Credibility of news having “poor” comprehensibility.

	TRUE	FALSE
Arabic	33	67
Chinese	27	73
Italian	0	100
Japanese	45	55
Vietnamese	23	77

Considering the impact of all the segmental and suprasegmental anomalous traits of the foreign speakers’ productions, two more tests based on artificially modified speech were planned, in order to perceptually confirm results obtained with natural speech.

3.3. Test 3

For the first step of the manipulation-based analysis, the attention was exclusively focused on tonal range and silent pauses. For this purpose, the two non-native voices used for the first test (advanced learners) were used in order to avoid that disfluencies or errors could somehow interfere with the experiment outcome. A control group composed by the test 1 native voices underwent a manipulation as well.

On the basis of the original native speech values, the tonal peaks and the silent pauses extension were artificially increased and decreased using WinPitch W7: the former underwent a rise or a flattening of about 50%, while the latter were shortened or lengthened by approximately 70%. These modifications were applied only to the 8 news that obtained polar values of credibility, i.e. maximum or minimum, in the first test.

A third perceptive test was then arranged and administered.

For all the modified voices, both native and non-native, data confirmed that less marked tonal variations and shorter silent pauses correlate with lower credibility values, while, by contrast, a restricted tonal range and more lasting silences increase listeners’ trust (Tables 8 and 9).

3.1. Test 4

The last step of the present study focused on anomalous segmental and suprasegmental traits, by using the transplantation technique. Each single piece of news was manipulated in two different steps, so that disfluencies and errors were removed, and the prosodic features of the native speaker’s utterances were transferred onto the same utterances produced by the non-native

speakers. The rhythmic-prosodic transplantation technique is based on the algorithm PSOLA implemented in Praat and illustrated in [24].

Table 5: Percentage values of credibility of original and modified speech: tonal range (TR).

	original TR	Synthesized speech		Δ
		decreased TR	increased TR	
% credibility	51		45	-6
	43	56		+13
	35	58		+23
	60		41	-19
	23	44		+21
	10	32		+22

Table 6: Percentage values of credibility of original and modified speech: silent pauses (SL).

	original SL	Synthesized speech		Δ
		Decreased SL	increased SL	
% credibility	51	42		-9
	52		57	+5
	43		58	+15
	35		38	+3
	60	50		-10
	10		17	+7

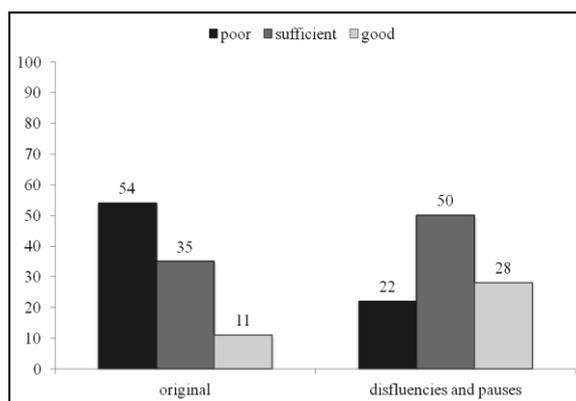
In the fourth perceptive test, in order to evaluate the impact of anomalous prosodic and segmental performances on the degree of the perceived foreign accent, listeners were asked to rate not only comprehensibility and credibility, but also foreign accent on a three-point scale (1= native speaker; 3= strong foreign accent).

For this experiment, the voices of the second test were used, but the native one has not been modified. The non-native speakers were all strongly foreign accented.

3.2. First step: removing disfluencies and cloning native pauses

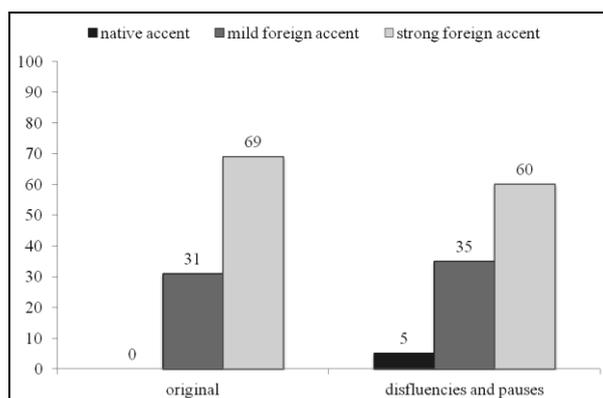
The perceptive test carried out after having removed all the disfluencies and cloned the native pauses, both in terms of position and duration, attested an increasing value of “sufficient” and “good” comprehensibility ratings for all foreigners (Fig. 4).

Figure 4: Comprehensibility: NNSs’ average values.



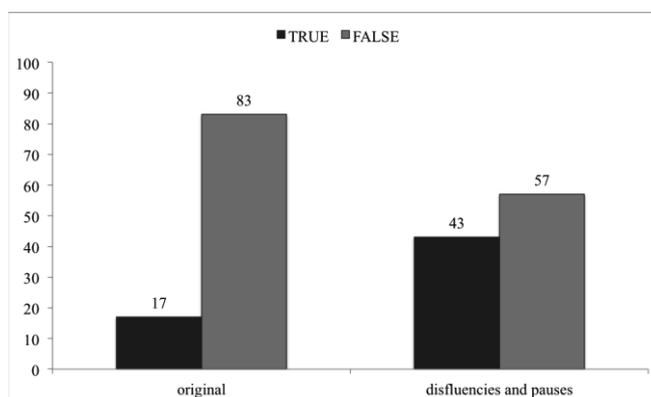
As for the foreign accent, almost all the listeners correctly recognized the NNSs (Fig. 5). The modifications carried out on NNSs’ utterances produced a slight decrease of the “strong foreign accent” ratings (from 69% to 60%). In addition, it is worth noting that the 5% of the listeners assumed to have heard a native voice.

Figure 5: Foreign accent: NNSs’ average values (%).



The removal of the disfluencies and the silences changes determined a significant increase in the level of news credibility (Fig. 6), taking the NNSs’ values to levels very close to those of the NS (true: 46%; false: 54%).

Figure 6: Credibility: NNSs’ average values (%).



3.2.1. Second step: errors removal

For the second step of the fourth phase of the study, data are limited to the A2 level NNSs, since in the other speakers' productions there were no particular segmental irregularities. Using WaveSurfer, segments perceptually detected as wrong were artificially modified or substituted by adequate micro-segments produced by the same speaker within the same utterance.

As regards the foreign accent, perceptive test data show that modifications of the errors produce a slight improvement. No significant variations occurred in terms either of comprehensibility or of credibility.

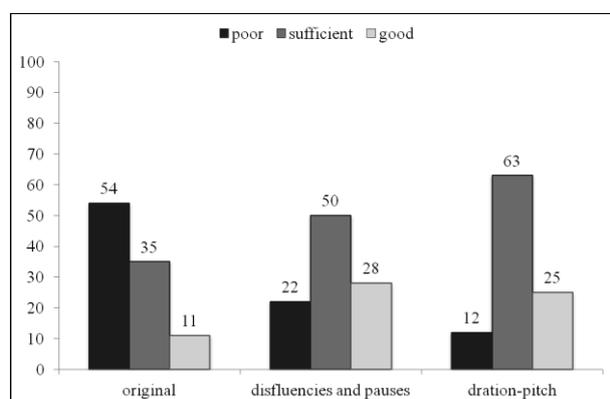
3.2.2. Third step: duration and pitch transplantation

The final part of the manipulation-based investigation consisted in the cloning of the segmental duration and the intonation contour from the utterances of NS to the utterances of NNSs.

Comparing the outcomes of this perceptive test with those concerning disfluencies removal and silences cloning, a further improvement of comprehensibility can be observed: the negative judgments decreased by 10% (from 22% to 12%) in favour of the "sufficient" ratings, while the percentage of "good" comprehensibility remained unchanged.

The comprehensibility gain achieved at the end of the whole transplantation process appears more noteworthy if compared to the results obtained by the original utterances: the "poor" comprehensibility reduces by 42% and the "good" comprehensibility better by 14% (Fig. 7).

Figure 7: Comprehensibility: NNSs' average values (%).



By contrast, credibility values (Fig. 9) did not undergo significant variations.

The most remarkable effects of the manipulation procedure are those relating to the foreign accent degree (Fig. 8): the modified NNSs' voices resulted in a lowering of 60% of the strong

foreign accent, leading, at the same time, the 30% of the listeners to maintain that they have heard an L1 Italian speaker.

Figure 8: Foreign accent: NNSs' average values (%).

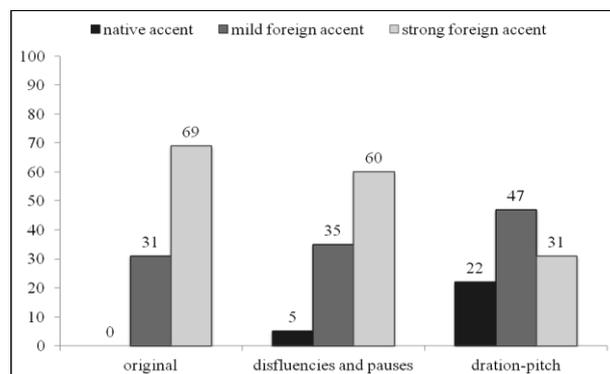
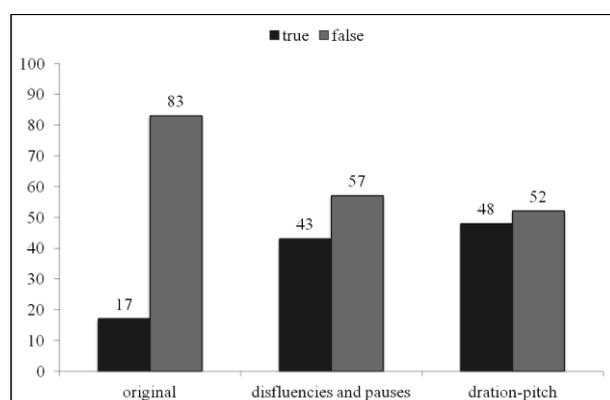


Figure 9: Credibility: NNSs' average values (%).



4. CONCLUSIONS

Going back to the question posed at the beginning of this study, whose aim was to assess the possible relation between foreign accent and credibility, findings suggest that credibility is closely related to the message comprehensibility, regardless of the speaker's accent (native accent, mild foreign accent, strong foreign accent).

Comprehensibility seems to be mostly affected by segmental and suprasegmental anomalies (disfluencies, wrong phones, inadequate silent pauses, incorrect tonal variations, etc.) which are typical of the productions of beginner and intermediate L2 speakers. Therefore, the lower the language skill level, the higher the percentage of prosodic and phonetic anomalies and, consequently, the lower the comprehensibility. A low level of understanding leads listeners to judge the message as not credible. Improving the speech quality better the understanding and, as a consequence, enhances the credibility.

In the case of a good level of comprehensibility, a speaker may improve his/her credibility by modulating intonation and silent

pauses: less varied tonal movements and longer silent pauses lead listeners to assign more easily a judgment of truth to the message heard.

Current findings support the theory of “processing fluency”, but at the same time indicate that there is only an indirect relationship between foreign accent and credibility, mediated by comprehensibility. A NNS having a strong foreign accent and an advanced level of L2 competence has the same chance of communication success of a NS, in terms of credibility.

A further and final phase of this research may involve the use of a corpus composed by read speech of disfluent native Italian speaker.

5. REFERENCES

- [1] Baroni, M. R. 1983. *Il linguaggio trasparente. Indagine psicolinguistica su chi parla e chi ascolta*. Bologna: Il Mulino.
- [2] Boersma, P., Weenink, D. 2012. *Praat: doing phonetics by computer*. Version 5.3.31, <http://www.praat.org/>
- [3] Boula de Mareuil, P., Vieru-Dimulescu, B. 2006. The contribution of prosody to the perception of foreign accent. *Phonetica* 63, 247–267.
- [4] Catricalà, M., Di Ferrante, L. 2010. Pregiudizi in movimento: un’inchiesta di matched guise in Italia e la questione dei diritti linguistici. In: Agresti, G., Bienkowski, F. (eds), *Les droits linguistiques: droit à la reconnaissance, droit à la formation*. Rome: Aracne, 158-178.
- [5] De Meo, A., Pettorino, M., Vitale, M. 2012. Non ti credo: i correlati acustici della credibilità in italiano L2. In: Bernini, G., Lavinio, C., Valentini, A., Voghera, M. (eds), *Atti dell’XI Congresso dell’Associazione Italiana di Linguistica Applicata. Competenze e formazione linguistiche. In memoria di Monica Berretta*. Perugia: Guerra Edizioni, pp. 229-248.
- [6] De Meo, A., Vitale M., Pettorino, M., Martin P., 2011. Acoustic-perceptual credibility correlates of news reading by native and Chinese speakers of Italian. In: Lee, W.-S., Zee E. (eds), *Proceedings of the 17th International Congress of Phonetic Sciences*, City University of Hong Kong, Hong Kong, Cina, 1366-1369.
- [7] Galli de’ Paratesi, N. 1984. *Lingua toscana in bocca ambrosiana. Tendenze verso l’italiano standard: un’inchiesta sociolinguistica*. Bologna: Il Mulino.
- [8] Gluszek, A., Dovidio, J.F. 2010. The way they speak: Stigma of non-native accents in communication. *Personality and Social Psychology Review* 14, 214-237.
- [9] Lev-Ari, S., Keysar, B. 2010. Why don’t we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology* 46, 1093-1096.
- [10] Metzger, M.J., Flanagin, A.J., Eyal, K., Lemus, D.R., McCann, R.M. 2003. Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. *Communication Yearbook* 27, 293-335.
- [11] Munro, M.J., Derwing, T.M. 2001. Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies of Second Language Acquisition* 23, 451-468.
- [12] Oppenheimer, D.M. 2008. The secret life of fluency. *Trends in Cognitive Science* 12, 237-241.
- [13] Pettorino, M., De Meo, A., Pellegrino, E., Salvati, L., Vitale, M., 2011. Accento straniero e credibilità del messaggio: un’analisi acustico-percettiva. In: Gili Fivela, B., Stella, A., Garrapa, L., Grimaldi, M. (eds), *Contesto comunicativo e variabilità nella produzione e percezione della lingua, Atti del 7° Convegno Nazionale dell’Associazione Italiana di Scienze della Voce (AISV 2011)*. Roma: Bulzoni editore, CD.
- [14] Piske, T., MacKay, I.R.A., Flege, J.E. 2001. Factors affecting degree of foreign accent in an L2: A Review. *Journal of Phonetics* 29, 191-215.
- [15] Reber, R., Schwarz, N. 1999. Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition* 8, 338-342.
- [16] Self C. S. 1996. Credibility. In: Salwen, M., Stacks, D. (eds.), *An integrated approach to communication theory and research*. Mahwah, NJ: Lawrence Erlbaum Associates, 421-441.
- [17] Thorne, S. 2005. Accent pride and prejudice: Are speakers of stigmatized variants really less loyal? *Journal of Quantitative Linguistics* 12, 151-166.
- [18] Trofimovich, P., Baker, W. 2006. Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition* 28, 1-30.
- [19] Tsalikis, J, DeShields, O. W. Jr., LaTour M. S. 1991. The Role of Accent on the Credibility and Effectiveness of the Salesperson. *The Journal of Personal Selling and Sales Management* 11/1, pp. 31-41.
- [20] Volkart-Rey, R. 1990. *Atteggiamenti linguistici e stratificazione sociale. La percezione dello status sociale attraverso la pronuncia. Indagine empirica a Catania e a Roma*. Roma: Bonacci.
- [21] Wathen, C. N., Burkell, J. 2002. Believe it or not: factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology* 53 (2), 134-144..
- [22] Wave Surfer 1.8.8. <http://www.speech.kth.se/wavesurfer>.
- [23] WinPitch. <http://www.winpitch.com/>
- [24] Yoon, K. 2007. Imposing Native Speakers’ Prosody on Non-native Speakers’ Utterances: The Technique of Cloning Prosody. *Journal of the Modern British & American Language & Literature* 25(4), 197-215.