# Identifying foreign speakers with an unfamiliar accent or in an unfamiliar language

*Andrea Weber*[1] *and Katja Pöllmann*[1,2]

[1]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[2]Universität Bamberg, Bamberg, Germany

andrea.weber@mpi.nl, katja.poellmann@mpi.nl

**ABSTRACT**

For an accent rating study, we recorded 12 English and Dutch short sentences with 6 native speakers of English and with 6 native speakers of Dutch; each speaker read all sentences, thus performing in both her native language and in her second language. The sentences were presented to 3 listener groups who all knew English but varied in their knowledge of Dutch: Dutch listeners, German listeners who knew Dutch, and German listeners who did not know Dutch. Listeners rated all sentences on a scale from 1 (strong foreign accent) to 9 (definitely native speaker). For the English sentences, we found that even German listeners who were not familiar with Dutch were as good as Dutch listeners themselves in identifying Dutch non-native speakers. Thus, familiarity with the language of the accent did not necessarily influence performance. For the Dutch sentences, we found that familiarity with the target language had a noticeably greater influence on responses. That is, German listeners without any knowledge of Dutch were less good at identifying non-native speakers of Dutch than German listeners with Dutch knowledge were. Interestingly, German listeners without any knowledge of Dutch were still better than chance at identifying non-native English speakers in Dutch sentences. The results suggest that familiarity with the target language may play a bigger role in accent detection than familiarity with the accent language, but in the absence of any knowledge of the target language listeners can still reliably detect foreign speakers.

**Keywords:** L2 speech, accent detection, L2 listener, L1 listener, familiarity.

## 1. INTRODUCTION

Native-like pronunciation in a second (L2) language is often seen as the most prominent and persistent difficulty for adult L2 learners. This can be problematic for successful communication because native (L1) listeners often have a negative attitude towards speakers with a foreign accent (e.g., Munro et al. 2006), and are known to be highly sensitive to foreignness in speech. For example, it has been shown that L1 listeners can detect L2 speakers in short stretches of speech (e.g., Flege 1984), in filtered speech (e.g., Munro 1995), or when the traces of accentedness are only minimal (e.g., Munro et al., 2003). An often used task for revealing sensitivity to foreignness in L2 speech is accent rating. In accent-rating studies, words or sentences are recorded with L1 and L2 speakers, and listeners are asked to rate on a scale how native-like the speech sounds; L1 listeners then reliably judge L1 speakers to be more native-like than L2 speakers. A debated issue in accent-rating studies is thereby whether listeners' responses indicate something about the speech itself because they are influenced by its acoustic and phonological properties, or whether they indicate something about the listener and therefore vary with the listeners' language experience, for example. Our approach to this issue was to focus on accent ratings from L2 listeners, a listener population for which there is comparatively little evidence for sensitivity to non-nativeness yet (but see Flege 1988; MacKay et al. 2006; Major 2007), and to vary their familiarity with the language of the accent but also their familiarity with the target language (i.e., the language they are asked to rate). In doing so we want to establish whether familiarity with either the target language or the accent language are a prerequisite for detecting non-nativeness in speech, and whether the degree to which familiarity with the target language and the accent language are beneficial, varies.

For our accent rating study, we used short English and Dutch sentences that had been produced by either native or non-native speakers. The L2 speakers for the English sentence were Dutch L1 speakers, and the L2 speakers for the Dutch sentences were English L1 speakers. While our listeners were all highly proficient in English, they varied in their knowledge of Dutch: they were either native listeners of Dutch, German listeners who knew Dutch, or German listeners who did not know any Dutch. German listeners with Dutch knowledge were students living in the Netherlands, immersed in a Dutch-speaking environment. German listeners without Dutch knowledge were students in the south of Germany (i.e., not close to the Dutch border) who have never learned Dutch or been to the Netherlands. Of particular interest was how well German listeners without any knowledge of Dutch could detect Dutch L2 speakers in English sentences, and whether the same listener group could identify English L2 speakers in Dutch sentences at all.

## 2. ACCENT RATING

### 2.1. Methods

#### *2.1.1. Participants*

Twenty-four native speakers of Dutch, with an average age of 21, took part in the experiment for a small monetary compensation. All 24 Dutch participants were students at the Radboud University in Nijmegen, the Netherlands, and they had received between 5 and 13 years of formal training in English (average of 8 years). In addition, 21 native speakers of German (average age of 23), also all students at the Radboud University, participated. At the time of testing, they had been living in the Netherlands for 2 years on average (ranging from 1 to 5). After their arrival in the Netherlands, they had received an intensive course in Dutch, and judged their level of proficiency in Dutch to be highly advanced, regularly speaking and listening to Dutch for more than 10 hours per week. They had received formal training in English for 9 years on average. The third group of participants consisted of 20 German listeners without any knowledge of Dutch. These participants were all students at the University of Bamberg in Germany (average age of 24). None of them had studied Dutch or had ever been to the Netherlands. All 20 students from Bamberg declared in a questionnaire to have no knowledge of Dutch, never to speak or listen to Dutch, and 17 of the 20 declared not to be familiar with Dutch accented-English at all (the other 3 stated to be a little bit familiar with it).

#### *2.1.2. Materials and Procedure*

Six short English declarative sentences and six short Dutch declarative sentences were chosen as materials (see Table 1). The English sentences were based upon the Bamford-Kowal-Bench sentences (Bench et al. 1979) and the Dutch sentences upon the Plomp and Mimpen (1979) sentences. Care was taken that the English and Dutch sentences were comparable in number of words and syllables. Some of the English sentences contained sounds that were judged to be typical sounds of American English (e.g., /ɝ/ in *stirs*), and others were less obviously marked. Dutch- and German-accented English share some markers of foreignness (e.g., in both accents, English /æ/ is usually produced as /ɛ/; see Swan, & Smith, 2001), and we therefore avoided these known overlapping markers to a large extent in the English sentences. Some of the Dutch sentences contained sounds that are typical of Dutch, and for which it is well known that many learners of Dutch have difficulties producing them correctly (e.g., /œy/ in *duinen*). Furthermore, some of the Dutch sentences had sounds that when produced by American learners of Dutch can be typical markers of an American accent (e.g., the /ɾ/ in *prijzen* incorrectly produced as /ɹ/).

**Table 1:** English sentences and Dutch sentences (English translation in brackets) used in the accent rating study.

| English sentences | Dutch sentences |
|---|---|
| 1. He cut his finger. | 1 Alle prijzen waren verhoogd.<br>(All prices were increased.) |
| 2. He found his brother. | 2. Het vliegtuig vertrekt over een uur.<br>(The airplane leaves in one hour.) |
| 3. Somebody took the money. | 3. De schrijver is goed.<br>(The composer is good.) |
| 4. The mother stirs the tea. | 4. Het meisje stond te wachten.<br>(The girl was waiting.) |
| 5. The girl has a picture book. | 5. De bus is niet op tijd.<br>(The bus is not on time.) |
| 6. The pond water is dirty. | 6. Het was leuk in de duinen.<br>(It was nice in the dunes.) |

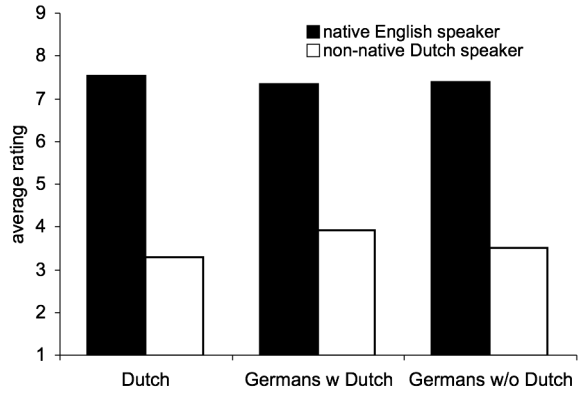
All 12 sentences were recorded by 6 native speakers of American English (2 male and 4 female, average age of 39, ranging from 27 to 68) and by 6 native speakers of Dutch (2 male and 4 female, average age of 46, ranging from to 28 to 71). At the time of the recording, the American speakers had been living in the Netherlands for 3 years on average. While the American speakers judged their proficiency in Dutch to be lower than the Dutch speakers estimated their proficiency in English, all speakers could produce the sentences in their L2 at a normal speech rate and without any noticeable disfluencies. Two independent native listeners furthermore ascertained that none of the speakers was completely accent-free. Sentences were recorded in a quiet room on a digital recorder at 44.1 kHz sampling rate with 16-bit resolution and were later transferred to a computer. Dutch sentences were on average 1.75 s long when produced by the Dutch speakers and 1.77 s when produced by the American speakers; English sentences were on average 1.39 s long when produced by the American speakers and 1.42 s when produced by the Dutch speakers.

The 144 sentences (12 sentences x 12 speakers) were presented in two blocks. One block contained all English sentences spoken by both English and Dutch speakers, and the other block contained all Dutch sentences spoken by both Dutch and English speakers. The order of blocks was counterbalanced across participants, and each participant heard the sentences in a different randomized order with the restriction that no sentence or speaker was heard twice in a row. Listeners were instructed to rate the sentences on a 9-point scale ranging from 1 (strong foreign accent) to 9 (definitely native speaker). Sentences were only played once, and participants were encouraged to use the full range of the scale for their ratings. The experiment started with two practice sentences, one English and one Dutch sentence, and it took approximately 25 minutes to complete the experiment.

### 2.2. Results

Figure 1 shows the average ratings for the English sentences by the 3 listener groups. As can immediately be seen in Figure 1, all 3 listener groups rated the English sentences significantly higher (i.e., more native-like) when the sentences had been produced by English L1 speakers than when the same sentences had been produced by Dutch L2 speakers.

**Figure 1:** Average ratings for English sentences by Dutch listeners, German listeners with Dutch knowledge, and German listeners without Dutch knowledge; 1 = strong foreign accent, 9 = definitely native speaker.




Analyses of Variance (ANOVAs) on the English sentences with the factors *listener group* (with the 3 levels 'Dutch', 'German with Dutch knowledge', and 'German without Dutch knowledge') and *speaker* (with the 2 levels 'English' and 'Dutch') showed a weak effect of *listener group* ($F_1[2,59] = 2.62$, $p > .08$; $F_2[2,10] = 5.13$, $p < .03$), a strong effect of *speaker* ($F_1[1,59] = 806.43$, $p < .001$; $F_2[1,5] = 233.57$, $p < .001$), and a significant interaction ($F_1[2,59] = 5.95$, $p < .005$; $F_2[2,10] = 20.06$, $p < .001$). Subsequent pairwise comparisons revealed that the interaction was mainly driven by Dutch listeners being better at differentiating between native English and non-native Dutch speakers than German listeners with Dutch knowledge were. Even though numerically this difference was rather small (Dutch listeners: 7.5 for L1 speaker and 3.2 for L2 speaker; Germans with Dutch knowledge: 7.3 for L1 speaker and 3.9 for L2 speaker), the interaction was significant. While this interaction suggests that sharing L1 language background with the L2 speaker can make accent detection easier, the lack of such a strong interaction between Dutch listeners and German listeners without Dutch knowledge, on the other hand, suggests that the advantage of sharing L1 language background with the L2 speaker does not always help. While it seems somewhat counterintuitive that familiarity with the accent language did not seem to enhance the ability to detect foreignness for L2 listeners, we want to note that previous studies also varied on whether they found a correlation between familiarity with the accent and ratings or not (e.g., Major, 2007; Munro et al., accepted).

**Figure 2:** Average ratings for Dutch sentences by Dutch listeners, German listeners with Dutch knowledge, and German listeners without Dutch knowledge; 1 = strong foreign accent, 9 = definitely native speaker.

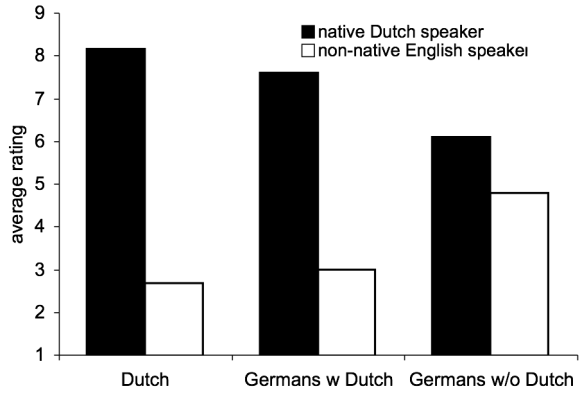



Figure 2 shows the average ratings for the Dutch sentences. As can be seen, Dutch listeners made again the clearest distinction between native Dutch speakers and non-native English speakers: they rated Dutch sentences significantly higher (i.e., more native-like) when the sentences had been produced by native Dutch

speakers than when the same sentences had been produced by non-native English speakers. This is not surprising, given that Dutch listeners were rating sentences in their native language. In overall ANOVAs, this time no significant effect of *listener group* ($F_1[2,59] = 1.71$, $p > .1$; $F_2[2,10] = 1.51$, $p > .2$), but a significant effect of *speaker* ($F_1[1,59] = 1324.24$, $p < .001$; $F_2[2,10] = 198.44$, $p < .001$), and a significant interaction ($F_1[2,59] = 136.18$, $p < .001$; $F_2[2,10] = 162.91$, $p < .001$) was found. Further pairwise comparisons showed highly significant interactions of *listener group* and *speaker* between each of the 3 listener groups. Thus, German listeners with Dutch knowledge rated native Dutch speakers lower and non-native English speakers higher than Dutch listeners did. While this difference could simply reflect differences between ratings from L2 and L1 listeners, a comparison with the second group of German L2 listeners showed that knowledge of the target language Dutch further modulated the rating patterns: German listeners without any knowledge of Dutch rated native Dutch speakers still lower and non-native English speakers still higher than German listeners with Dutch knowledge did. This suggests that familiarity with the target language helps L2 listeners to detect foreignness in speech. But as the difference in ratings by German listeners without any knowledge of Dutch further suggests, familiarity with the target language is not a requirement to detect foreignness: even though German listeners from Bamberg did neither speak nor understand Dutch they were not at chance when they had to detect foreignness in the Dutch sentences as a significant *speaker* effect showed ($F_1[1,18] = 54.24$, $p < .001$; $F_2[1,5] = 21.58$, $p < .007$).

## 3. DISCUSSION

For the English sentences, we found that Dutch listeners, as well as German listeners with and without knowledge of Dutch could distinguish between English and Dutch speakers very well; that is, they all judged sentences from non-native Dutch speakers to be more accented than sentences from native English speakers. Familiarity with the accent language Dutch, however, did not increase German L2 listeners' ability to detect Dutch L2 speakers. In fact, German listeners with Dutch knowledge were slightly less good at detecting L2 speakers than German listeners without Dutch knowledge were. For the Dutch sentences, on the other hand, we found that familiarity with the target language did influence responses strongly. German L2 listeners without any knowledge of Dutch were less good at identifying non-native English speakers than German listeners with Dutch knowledge were (and not surprisingly they were also less good than Dutch L1 listeners were). But importantly, German listeners without any knowledge of Dutch were still better than chance at identifying non-native English speakers in Dutch.

Presumably, native listeners can identify foreign speakers by being able to perceive whether the speech signal deviates from native pronunciation or not. Similarly, proficient L2 listeners may be able to base their judgments on whether or not the speech signal corresponds to the pronunciation norms of the listeners' L2. L2 listeners who are not familiar with the language being spoken, however, cannot base their judgments on recognizing how much the pronunciation corresponds to the target language norms because they do not know what the correct pronunciation of a sound or word in the target language is. How then could German listeners without any Dutch knowledge detect non-native speakers in Dutch?

One possibility is that sentence durations influenced their accent ratings. Non-native speakers often speak more slowly than native speakers, and previous studies have shown that listeners rate more slowly produced speech to be more accented than faster speech (e.g., Munro & Derwing, 1998). Note that sentence durations in the present study were not on average longer when sentences had been produced by non-native speakers than when they had been produced by native speakers. Furthermore, we found no significant correlation between ratings and sentence length (p-values for all correlations larger than .3). But the possibility remains, of course, that single words or syllables within the sentences were produced with a slower speaking rate in the non-native recordings and that German L2 listeners picked up on this cue.

A second possibility is that we underestimated our listeners' knowledge of Dutch. All German listeners from Bamberg stated in a questionnaire that they did not understand or speak any Dutch, but they probably nevertheless had some idea of what Dutch sounds like. If we would ask them whether a spoken Dutch sentence is Dutch or Chinese, we assume that they could reliably decide that the sentence is Dutch. They

may also have heard Dutch-accented German before. But then they still should not know whether the sounds and words in the Dutch sentence were produced correctly or bear traces of foreignness.

A third option for how German listeners without Dutch knowledge detected L2 speakers in Dutch sentences has to do with familiarity with the accent language English (either from native English or English-accented German). Our German listeners were highly proficient in English, and they may have based their ratings on recognizing whether the speech signal contained segmental and suprasegmental characteristics of the accent language English or not (still not knowing whether the same characteristics might be Dutch too). We tried to look into this possibility by analysing a subset of our sentences that we judged to be relatively free of segmental accent markers that are typical of an American English accent (e.g., the sentences contained no /r/; Dutch sentences 4-6 in Table 1). Indeed, we found that the difference in average ratings between native and non-native speakers became somewhat less pronounced (while ratings for native Dutch speakers hardly changed, ratings for non-native English speakers rose from 4.8 to 5.3). The remaining difference in ratings was, however, still significant. Thus, perceiving segmental characteristics of the accent language alone probably cannot explain how German listeners without Dutch knowledge detected foreignness in Dutch sentences. This leaves the possibility that listeners responded to something more general in the speech of non-native speakers. Voice quality, for example, has been suggested as a potential marker of nonnativeness, though its role in L2 production has not been thoroughly investigated yet (e.g., Esling 2000). Articulatory effort and carefulness are other potential markers. While with the present study we cannot uncover such possible language-independent, general markers and our results probably reflect a combination of all the above mentioned possibilities, the results of the German listeners make the existence of such general markers more likely. With respect to the question of whether the speech or the listener influence accent ratings, the present findings suggest that it is both. Linguistic experience improves the ability to detect foreignness in speech (speaking for listener effects), but the ability to detect foreignness without experience further suggests a role for the speech signal itself.

## 4. REFERENCES


Bench J., Kowal A., Bamford J. 1979. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology* 13. 108-112.

Esling, J. H. 2000. Crosslinguistic aspects of voice quality. In R. D. Kent, M. J. Ball (Eds.). *Voice Quality Measurement* (pp. 25-35), Singular Publishing Group, San Diego.

Flege J. E. 1984. The detection of French accent by American listeners. *Journal of the Acoustical Society of America* 76. 692-707.

Flege J. E. 1988. Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America* 84. 70-79.

MacKay I. R. A., Flege J. E., Imai S. 2006. Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics* 27. 157-183.

Major R. C. 2007. Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition* 29, 539–556.

Munro M. J. 1995. Nonsegmental factors in foreign accent: ratings of filtered speech. *Studies in Second Language Acquisition* 17. 17-34.

Munro M. J., Derwing T. M. 1998. The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning* 48. 159-182.

Munro M. J., Derwing T. M., Burgess C. S. 2003. The detection of foreign accent in backwards speech. In M-J. Sole, D Recasens, J. Romero (Eds.). *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 535-538). Barcelona, Spain: University Autonoma de Barcelona.

Munro M. J., Derwing T. M., Morton S. L. 2006. The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition* 28. 111-131.

Munro M. J., Derwing T. M., Sato K. 2006. Salient accents, covert attitudes: Consciousness-raising for pre-service second language teachers. *Prospect* 21. 67-79.

Munro M. J., Derwing, T. M., Burgess C. S. accepted. Detection of nonnative speaker status from content-masked speech. *Speech Communication*.

Plomp R., Mimpen A. M. 1979. Speech-reception threshold for sentences as a function of age and noise level. *Journal of the Acoustical Society of America* 66. 1333-1342.

Swan M., Smith B. (Eds.). 2001. *Learner English. A teacher's guide to interference and other problems*. Cambridge: Cambridge University Press.