PROSODIC ACCOMMODATION BY FRENCH SPEAKERS TO A NON-NATIVE INTERLOCUTOR

Caroline L. Smith

University of New Mexico
caroline@unm.edu

ABSTRACT

On-line accommodation to an interlocutor is often cited as an explanation for phonetic variation. Prosodic evidence for speakers' accommodation was investigated in a task that was expected to favor modification: giving directions to a non-native interlocutor, compared to the same task with a native interlocutor. Ten native speakers of French were recorded in spontaneous conversation in the two conditions. With the non-native, they used a significantly greater F0 range, and segmental modifications compatible with a more emphatic speech style, but did not modify speech rate or utterance duration. These results suggest that accommodations can include both language-specific and universal properties, and that speakers can selectively implement different ways of accommodating.

Keywords: prosody, foreigner talk, speech rate, vowel devoicing, French

1. INTRODUCTION

Lindblom's [8] "Hyper- & Hypo- speech" (H&H) theory claims that speaker accommodation to the listener and the situation is a major reason for the variability in the speech signal. An example of accommodation is speaking more loudly when in a noisy environment (the Lombard effect).

The question asked here is: To what extent do speakers make on-line, fine-grained modifications for the benefit of the listener? Numerous studies support the intuition that speakers modify their speech depending on the situation, but the evidence is much more mixed as to whether such modifications are intended to benefit the listener, are done in a way that would in fact benefit the listener, or whether such modifications are instead driven primarily by the speaker's own needs and perceptions [1, 12].

The situation being studied here is accommodation by a native speaker to a non-native interlocutor (NNI). Native speakers' speech to NNIs, known as "foreigner talk", is described as having exaggerated intonation, louder and slower speech, but reports vary as to whether speakers use more or less phonological simplification to NNIs than to native interlocutors (NIs) [7]. There has been remarkably little systematic study of the acoustic characteristics of speech by native speakers to NNIs. Of the five recent studies identified, two compared speakers' productions to imaginary NIs and NNIs [2, 10], a third compared radio broadcasts for learners of English with broadcasts for a general English-speaking audience [9], and only two compared speech produced to physically present NIs and NNIs [11, 14].

These studies give conflicting evidence for the types of phonetic accommodation speakers produce to NNIs. Speech rate has been found to be slower to NNIs [2, 9, 10], although vowels in monosyllabic words were not significantly longer in [14]. Utterance duration is reported as either longer [10] or shorter [9] to NNIs. Speakers take longer pauses when speaking to NNIs [2, 9] and use a more expanded vowel space [11, 14], which is typical of clear speech. One study [10] found speakers used a greater range of F0 to NNIs (presumably what anecdotal reports of exaggerated intonation are referring to), but another study [2] did not find a significant difference in F0 range, and mean F0 did not differ in [14]. An expanded F0 range enhances intelligibility [3], possibly by providing clearer intonational contrasts.

The study here focuses on phonetic accommodation to an interlocutor who is reasonably fluent, but clearly non-native. If on-line accommodation is truly pervasive, evidence of it should be found even in speech to an interlocutor who has a good command of the language.

2. METHOD

Recordings were made of ten native French speakers, who each participated in a pair of telephone conversations. Each conversation required the speaker to perform a map task, giving directions for travel by Métro to various locations

around Paris. Half of the speakers spoke first with a native French interlocutor (NI), then with a non-native interlocutor (NNI) who also spoke in French. The other half of the speakers spoke with the non-native first. Each interlocutor asked to travel to a different set of destinations.

The native French speakers were all female undergraduates in Paris. They were naive to the purpose of the study. The interlocutors were colleagues who were aware of the purpose. The native French interlocutor identified herself as a business person visiting Paris; the non-native interlocutor identified himself as an American researcher in Paris for a conference. The native French interlocutor judged that all the speakers were using a dialect typical of the Paris region. The interlocutors were not known to the speakers, nor did they see them.

The conversations were recorded digitally in a sound-treated room at the Université de Paris III and analyzed by inspection of the waveform and spectrogram in Praat.

2.1. Timing measures

The durations of these conversations ranged from 4.3 to 11.4 minutes (mean 7.6). For all but one of the ten speakers, the conversation with the non-native interlocutor was longer than the conversation with the native French interlocutor.

The conversations were transcribed and divided into utterances demarcated by any silent pause or laughter longer than 150 ms. Only those "utterances" were retained that contained words or speech sounds with lexical meaning. That is, a noise such as *euh*… indicating hesitation was not treated as an utterance, but *ah* indicating agreement was retained. In order to test whether speakers used shorter utterances to the NNI, the average utterance duration was calculated for each speaker by summing the duration of all speech during utterances and dividing by the number of utterances. Note that this excludes pauses.

A second measure of timing, words per minute, was predicted to be lower for speech to the NNI. It was estimated for each conversation by taking the total number of words produced by the speaker, and dividing by the total duration of all of her utterances. The total number of words was determined from the transcribed utterances, with a word being a unit demarcated by a space in conventional orthography (so *qu'est-ce* is one word). This is described as an estimate because the nature of spontaneous speech makes an accurate word count difficult: here all partial words and hesitation sounds produced as part of a meaningful utterance are counted as words.

2.2. F0 measures

Most predictions about intonation in foreigner talk suggest that speakers use an expanded F0 range to NNIs, so increased range rather than mean value was chosen as the best test of this hypothesis. In order to measure the central tendency of speakers' F0 usage, and avoid undue influence from outlying regions of extreme values or errors from the pitch tracker, F0 range was calculated as the difference between the third quartile (75%) and first quartile (25%) of F0 values for each speaker.

2.3. French prosodic markers

This study uses two "prosodic markers" to investigate whether speakers modified their pronunciation by using these markers more or less frequently when speaking to the NNI. These are segmental modifications that are conditioned by prosodic factors.

The first, vowel devoicing, occurs in the common pronunciation of "oui" as [w̥i] where the final vowel is close to being a voiceless palatal fricative. Mid, low and nasalized vowels can also devoice, with a lower-amplitude voiceless portion that maintains the vowel's formant structure. Vowel devoicing is strongly associated with phrase-final position [13]. Although fairly common in contemporary speech, devoicing is still somewhat marked stylistically, occurring more often in reading than in conversation [5].

Vowel devoicing was identified from inspection of the spectrogram: 221 cases of devoicing in utterance-final position were identified, and 17 non-utterance-final. Further analysis was limited to the utterance-final cases. The frequency of devoicing was calculated as the percentage of utterances that ended with devoiced vowels.

The second prosodic marker, schwa addition, is identified as when a schwa is added to the end of a word, not as a separate hesitation sound, regardless of the presence or absence of orthographic *e* [4]. These "supporting" schwas can occur with words ending in any sound, and seem to be most frequent when the speaker is emphatic or annoyed [6].

Schwas were identified from inspection of the spectrogram. The frequency of their occurrence was calculated in "schwas per minute", by taking

the total number of added schwas produced by each speaker and dividing by the total duration of speech in utterances. This measure, rather than a per utterance measure, allows for the fact that some utterances contained multiple schwas.

2.4. Statistical tests

When data from all speakers was combined, paired t-tests were run in Microsoft Excel using the mean value for each speaker, with a significance level of .05. Tests of data from individual speakers used all the values for each utterance, and t-tests assuming unequal variance were run.

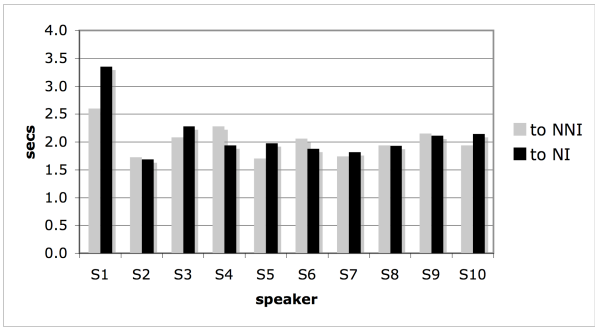
3. RESULTS

3.1. Timing measures

3.1.1. Average utterance duration

For all speakers together, there was no significant difference in the mean duration of utterances produced to the two interlocutors. The mean duration of utterances to the NNI was 2.024 s, to the NI, 2.112 s (t(9)=0.936, one-tail *ns*).

Figure 1: Mean utterance duration (s).

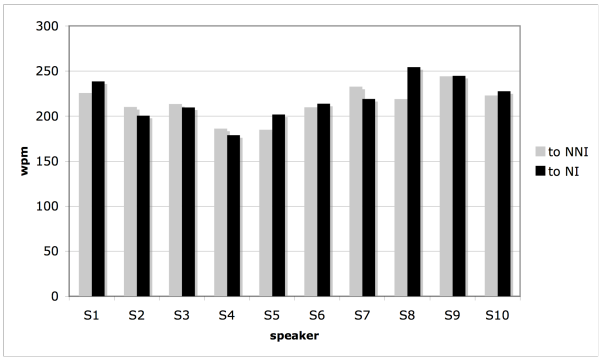


The exception to the overall lack of difference was speaker 1, who produced longer utterances to the NI (3.354 s vs. 2.603 s, t(129)=2.44, one tail p<.01). All other speakers showed no significant difference when tested individually.

3.1.2. Speech rate (words per minute)

The speakers did not alter their speech rate significantly when speaking to the NNI. The mean rate was 214 words per minute (wpm) to the NNI and 217 to the NI (t(9)=0.713, one-tail *ns*). Four speakers actually spoke more rapidly to the NNI, which is surprising as his speech rate, although not measured, was audibly far slower than that of the speakers or the NI. The largest difference for an individual speaker was S8, who spoke 35 wpm faster to the NI.

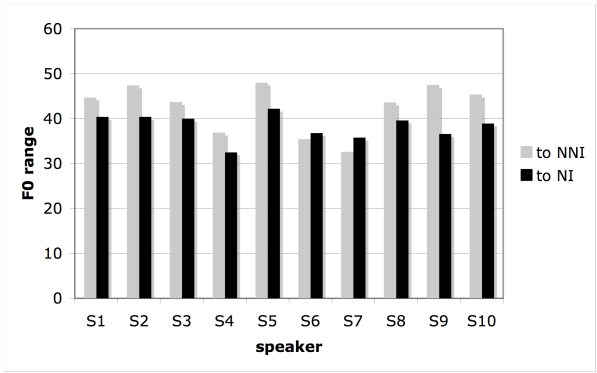
Figure 2: Speech rate in words per minute.



3.2. F0 range

Measured as the difference between the 75% quantile and the 25% quantile, the speakers' F0 range expanded when speaking to the NNI (mean of 42.5 Hz vs. 38.3 Hz to the NI, t(9)=3.286, one-tail p<.01). This expansion is consistent with the idea that speakers exaggerate intonation to NNIs.

Figure 3: F0 Range (75% quantile – 25% quantile), in Hz.

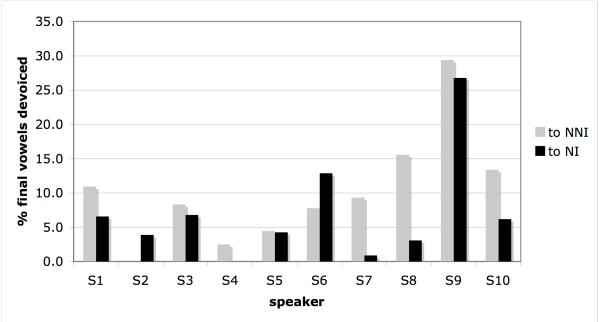


3.3. Prosodic markers

3.3.1. Utterance-final vowel devoicing

Individual speakers varied considerably as to the frequency with which they devoiced vowels, but overall, they devoiced utterance-final vowels more often when speaking to the NNI (10% of final vowels) than to the NI (7%). However, this difference was not significant for all speakers combined (t(9)=1.78, two-tail *ns*). Analyzing each speaker individually using z-tests to compare their frequencies of devoicing to the two interlocutors, S7, S8 and S10 devoiced significantly more often to the NNI than to the NI. S2 and S6 devoiced non-significantly more often to the NI.

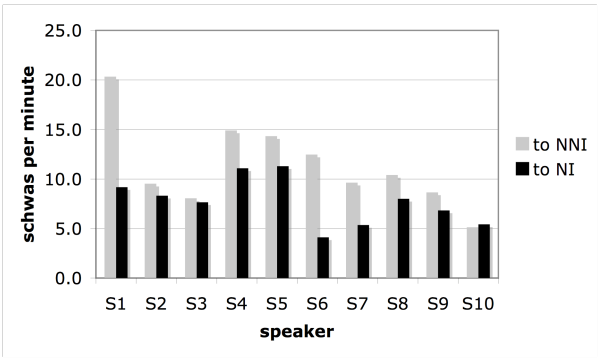
Figure 4: Percentage of utterances ending in a devoiced vowel.



3.3.2. Schwa addition

Speakers added schwas more often to the NNI, with an overall average of 11.3 per minute to the NNI, and 7.7 to the NI. This difference was significant ($t(9)=3.19$, two-tail $p=.011$). Only S10 did not conform to the overall pattern.

Figure 5: Schwas added per minute of speech.



4. DISCUSSION

When speaking to the NNI, speakers used a greater F0 range, more added schwas, and for some, more final vowel devoicing. The increased use of schwas is consistent with a more emphatic speaking style, which might reflect speakers' trying to be more comprehensible. But despite the NNI's slow rate of speech, the speakers did not slow down. Nor did they shorten their utterances, a form of simplification. The lack of change in rate is the most surprising finding, given that slower speech is a modification invariably mentioned in work on foreign language instruction, and is likely the modification most helpful to an NNI.

Vowel devoicing and added schwas are stylistic devices specific to French, and the increased use that some or all of the speakers made of these when speaking to the NNI shows that accommodations to interlocutors can include language-specific patterns as well as modifications often thought of as more universal, such as slower speech. These results suggest that "foreigner talk" is not a monolithic register. Accommodation may be possible along different dimensions in different languages, and speakers' choices about how to accommodate may not be maximally beneficial to the listener.

5. REFERENCES

[1] Bard, E., Anderson, A., Sotillo, C., Aylett, M., Doherty-Sneddon, G., Newlands, A. 2000. Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language* 42, 1-22.

[2] Biersack, S., Kempe, V., Knapton, L. 2005. Fine-tuning speech registers: a comparison of the prosodic features of child-directed and foreigner-directed speech. *Proc. of Eurospeech 2005* (Lisbon), 2401-2404.

[3] Bradlow, A., Toretta, G, Pisoni, D. 1996. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Comm* 20, 255-72.

[4] Fagyal, Zs. 2000. Le retour du *e* final en français parisien: changement phonétique conditionné par la prosodie. In: Englebert, A., et al. (eds.), *Vivacité et Diversité de la Variation Linguistique*. Tubingen: Max Niemeyer, 151-160.

[5] Fagyal, Zs., Moisset, C. 1999. Sound change and articulatory release: where and why are high vowels devoiced in Parisian French. *Proc. of the 14th ICPhS* (San Francisco), 309-312.

[6] Hansen, A. B. 1997. Le nouveau [ə] prépausal dans le français parlé à Paris. In Perrot, J. (ed.), *Polyphonie pour Iván Fónagy*. Paris: L'Harmattan, 173-198.

[7] Hatch, E. 1983. *Psycholinguistics: A Second-language Perspective*. Rowley, MA: Newbury House Publishers.

[8] Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In: Hardcastle, W., Marchal, A. (eds.), *Speech Production and Speech Modelling*. Dordrecht: Kluwer, 403-439.

[9] Osada, N. 2003. Analysis of pause occurrence in three kinds of modified speech: public address, caretaker talk, and foreigner talk. *Pan Pacific Association of Applied Linguistics* 7, 77-123.

[10] Papoušek, M., Hwang, S.-F. 1991. Tone and intonation in Mandarin babytalk to presyllabic infants: comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics* 12, 481-504.

[11] Poch, D., Dhainaut, C., Huet, K., Harmegnies, B. 2003. La adaptación del locutor al interlocutor: dinámica del sistema vocálico del español en función de las características de los participantes en la conversación. *Estudios de Lingüística Univ. de Alicante* 17, 519-530.

[12] Schober, M., Brennan, S. 2003. Processes of interactive spoken discourse: The role of the partner. In: Graesser, A., Gernsbacher, M., Goldman, S. (eds.), *Handbook of Discourse Processes*. Mahwah, NJ: Erlbaum, 123-165.

[13] Smith, C. 2003. Vowel devoicing in contemporary French. *J. of French Language Studies* 13, 177-194.

[14] Uther, M., Knoll, M., Burnham, D. 2007. Do you speak E-N-G-L-I-S-H? A comparison of foreigner-and infant-directed speech. *Speech Communication* 49, 2-7.