Intelligibility at a multilingual cocktail party: Effect of concurrent language knowledge

Aurore Gautreau 1, Michel Hoen 1,2, Fanny Mennier 1
1 Laboratoire Language, Cerveau et Cognition (L2C2), CNRS UMR5304, University of Lyon.
2 Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, CNRS UMR5292, INSERM U1028, University Lyon 1.
aurore.gautreau@gmail.com

Abstract
Our research aims at exploring the nature of the interferences that occur during the speech-in-speech situation. French target words were inserted in 2 types of backgrounds: (i) 4-talker babble spoken in various languages such as French, Italian or Irish, containing acoustic and linguistic information, (ii) fluctuating noise derived from each 4-talker babble signal, with only acoustic information. In Experiment 1, French native participants with no knowledge of Italian or Irish performed a lexical decision task. The comparison of performances obtained with the 2 types of backgrounds for each language revealed that acoustic and linguistic information from babble spoken in a known language to the participants (French) competed with the target words; whereas for babble produced in unknown languages (Italian and Irish) only acoustic information was involved. In Experiment 2, the experimental conditions were identical as in Experiment 1, except that French native participants speaking Italian as L2 and with no knowledge of Irish were recruited. The fact that Italian became an intelligible language to the participants led to acoustic and linguistic interferences from Italian babble.

Index Terms: multi-talkers babble, fluctuating noise, speech comprehension, acoustic interference, linguistic interference

1. Introduction
Comprehension of speech is more difficult with background speech (for example multiple talkers in a meeting room) than with any background without linguistic information (such as a traffic-jam), due to acoustic and linguistic interferences that occur between the target speech and the concurrent speech [1, 2]. Our research focused on the speech-in-speech situation in order to dissociate and evaluate the effects of acoustic and linguistic interferences from the background speech (babble).

Previous research on speech-in-speech comprehension demonstrated that the intelligibility of the target speech is lower with babble produced in the same language than in a different language. This effect has been highlighted in several studies examining the intelligibility of English target speech with various babble spoken either also in English or in a different language that was unknown to the native English-speaking participants, including Dutch [3], Spanish [4] and Mandarin [5]. The fact that the participants understand the linguistic content of the English babble led to linguistic interference, which thereby decreased their performances. Collectively, these studies revealed that the intelligibility of the babble influences the effect of the linguistic interference. More recently, Calandruccio et al. (2010) [6] proposed to English native participants an intelligibility task with English target sentences that were masked by various babble that varied in the amount of linguistic content. They were produced whether in English (as the target language), Mandarin (unknown language to the participants), or in English spoken by Mandarin native speakers whose production was evaluated has having high-, moderate- or low-intelligibility. Results indicated lower performances with the English babble, i.e. when the linguistic content of the babble was fully intelligible. As regards the English babble with high-, moderate- or low-intelligibility, the effect of the linguistic interference became more important as the babble became more intelligible.

In our study we further explored the composition of the linguistic interference to determine if different types of competitions, such as phonetic/phonological competitions can come into play in the speech-in-speech situation. To do so, French target words were masked with babble from an identical language as the target speech (French) or from 2 different languages, i.e. Italian and Irish that are unknown to the French native participants, and have varying distances from the target language. Indeed, Italian is much closer to French than Irish. Both French and Italian are Romance languages and syllable-timed languages; whereas Irish is a Goidelic language and a stress-timed language. Moreover, French is composed of 35 phonemes (21 consonants and 14 vowels); 60% of Italian phonemes are comparable to French phonemes while Irish only shares 18% of its phonemes with French (Maddieson et al., 2011 [7]). Two different groups of participants were recruited, in Experiment 1 French native participants with no knowledge of Italian or Irish, and in Experiment 2 French native participants speaking Italian as second language (L2) and with no knowledge of Irish.

First, we expected that French babble would lead to poorer performances and that the effect of the linguistic interference would be larger than that of babble spoken in the 2 unknown languages. Second, we observed if Italian and Irish babble would provoke equivalent performances and whether the distance criteria between the languages would have an influence. Furthermore, we explored how the knowledge of the concurrent language interferes with lexical decision.

We were also interested in dissociating acoustic and linguistic interferences from the background speech. So from babble generated in each language manipulated (French, Italian and Irish), we derived fluctuating noise for which linguistic information was removed [8]. In comparison to babble, fluctuating noise produces similar acoustic interference but no linguistic interference. Consequently, we expected that fluctuating noise should lead to better performances than babble. A comparison of performances obtained with these 2 types of backgrounds in each language manipulated would indicate whether the linguistic information contained in the babble competes with the French target words.

The experimental conditions (3 languages (French vs. Italian vs. Irish) * 2 backgrounds (babble vs. fluctuating noise)) were tested with a lexical decision task considered as
an online measure that allows capturing interferences that occur during the lexical access.

2. Experiments

In Experiments 1 and 2, all materials, methods and procedures were exactly identical.

2.1. Materials and methods

2.1.1. Participants

Thirty volunteers participated in Experiment 1 and 24 in Experiment 2. All were students aged 18 to 30 years, and none of the participants indicated having a known hearing loss or language disorder. All signed consent forms and were compensated for their participation. In Experiment 1, the participants were native speakers of French, and no knowledge of Italian or Irish. In Experiment 2, the participants were native of French with Italian as a L2 and had no knowledge of Irish. They had on average 8 years of education in Italian, and still receive Italian lessons at the university.

2.1.2. Four-talker babble

For each language used as a background (Irish, Italian and French), several female and male speakers were individually recorded in a soundproof booth, reading the same passages from the book, The Little Prince (in French: “Le Petit Prince”) by Antoine de Saint-Exupéry, in their native language. Recordings of 2 female and 2 male talkers who had the most natural speech style were selected for each language. Babble was composed of 4 talkers because prior research from our group [9] showed that this listening condition (1 target talker vs. 4 concurrent talkers) is optimal to dissociate the effects of the acoustic and linguistic interferences. All recordings were modified according to the following protocol: (i) removal of silences and pauses exceeding 500 ms, (ii) suppression of sentences containing pronunciation errors or proper nouns, and (iii) intensity calibration at 70 dB-A. Then, several sequences of 4 s were randomly extracted from each recording, and the 4-talker babble was generated by mixing one randomly chosen sequence of 4 s from each of the 4 talkers in one language. Twenty-six different 4-s-long sequences of 4-talker babble were ultimately created for each of the 3 languages.

2.1.3. Fluctuating noise

To obtain fluctuating noise with comparable energetic masking characteristics as those for the babble, we derived 78 fluctuating noise samples directly from the 78 samples of 4-talker babbles, according to the following protocol: using MATLAB© (R2010a, The MathWorks, Inc., Natick, Massachusetts, USA), we first computed the energy root mean square (rms) of the original sample and extracted its temporal envelope by applying a 60 Hz low-pass filter. Then, a fast Fourier transform (FFT) was used to extract the power spectrum and phase distribution of the original signals. Next, an inverse FFT was used to generate new signals in which the spectrum and phase distribution of the original signals were exactly identical. Finally, the rms powers of the original and new signals were equated.

2.1.4. Target words

Eighty-one French disyllabic words and 81 pseudo-words were recorded in a sound-proof booth by a female native speaker of French aged of 28 years old. Words were selected in a middle range of frequency of occurrence (ranging from 0.29 to 175.65 per million; mean = 17.16, SD = 30.43), according to the French database Lexique2 [10], in order to avoid extremely high- or low-frequency items that the participants typically overused or ignored. All pseudo-words respected French phonotactic rules, ex: draton.

2.1.5. Stimuli and word lists

The 162 stimuli consisted of 81 target words and 81 target pseudo-words, which were mixed with 4 s of background sound (i.e., babble or fluctuating noise) at an SNR of -5 dB. Target words were inserted 2.5 s from the start of a background sound. From the 81 target words, 3 were used as practice items. To ensure that each of the 78 remaining words was presented in each of the 6 conditions (2 backgrounds types (babble vs. fluctuating noise) * 3 languages (French vs. Italian vs. Irish)), 6 different experimental lists were generated. Across the lists, each of the 78 words was presented in all of the conditions. Each participant heard only one list, such that each target word was presented only once to each participant to avoid repetition effects. Within each list, the order of stimuli was randomized across participants to avoid presentation order effects. The same process was conducted with the pseudo-words.

2.1.6. Procedure

Participants were tested individually in a quiet room facing a computer monitor. Stimuli were delivered with E-prime, diotically via headphones (Sennheiser HD 448) at a comfortable sound level (output fixed at 65 dB SPL, as measured with an artificial ear). Participants were instructed to perform a lexical decision task on the target items that were inserted in the background sounds. Their task was to decide as quickly and accurately as possible whether the target item was a word by pressing 1 of 2 pre-selected keys on a computer keyboard. Prior to the testing phase, participants were given 12 practice items to accommodate themselves with the stimulus presentation mode and the target voice. The experiment lasted an average of 30 min.

2.2. Results

For each of the 2 experiments, an analysis of variance (ANOVA) was conducted on mean reaction times (RTs: time-interval in milliseconds between the onset of the target speech and the participants’ button press) that were measured for the correct responses to the target words in each experimental condition. Trials in which participants made mistakes (Experiment 1 : 34.1%; Experiment 2: 34.1%), provided no response during the allotted time of 4,500 ms (Experiment 1: 4.3%; Experiment 2: 9.4%), or had RTs lower than 300 ms (Experiment 1: 0.2%; Experiment 2: 0.3%), were not included in the analyses. Thus, RT was considered as the dependent variable and background (babble vs. fluctuating noise) and language (Irish vs. Italian vs. French) as the within subjects factors.

2.2.1. Experiment 1 – Italian unknown language

Results are shown in Figure 1. The analysis revealed a significant main effect of background (F(1,29)=8.92, p=.005). Mean RTs were longer with babble in the background (1,234 ms) than with fluctuating noise (1,167 ms). The results also revealed a significant main effect of language (F(2,58)=5.83, p<.005). In average, participants were faster when the background was Irish (1,146 ms) and slower when it was French (1,217 ms) and they were slowest when the
background was Italian (1,239 ms). The interaction between these 2 factors was not significant (F(2,46)=1.25, p=.29). Post-hoc comparisons with the HSD Tukey test showed a tendential difference between Irish and Italian babble (p=.07) and there was a significant difference between Irish and French babble (p<.05). Finally, the effect of background was present only for French (p=.02), with RTs significantly faster when the background was fluctuating noise compared with babble.

In Experiment 1, we therefore observed the expected effect of background, confirming that the listening situation was easier when the background was composed of acoustic information alone rather than acoustic and linguistic information together. Then, for French, i.e. the native language of the participants, the significant difference between fluctuating noise and babble suggested that acoustic and linguistic information contained in the French babble both slowed the lexical access of the French target words. For Irish and Italian, i.e. unknown languages to the participants, the lack of significant difference between the 2 kinds of backgrounds (fluctuating noise vs. babble) indicated that only acoustic information from the Italian and Irish babbles disrupted the lexical access of the target words.

Results from Experiment 1 also showed that the Italian babble (unknown language) led to RTs as long as the French babble (native language) despite the fact that the nature of their interferences was different.

In a second experiment, we further explored the effect observed between the Italian and French babbles by presenting the same experimental conditions as in Experiment 1 to French native participants speaking Italian as a L2. Now that Italian is an intelligible language to these participants, it was question to determine if linguistic information from Italian babble will compete with the French target words.

2.2.2. Experiment 2 – Italian as L2

Results are shown in Figure 2. They indicated a significant main effect of background (F(1,23)=17.43, p<.001), with longer mean RTs when background was composed of babble (1,347 ms) rather than fluctuating noise (1,214 ms). The analysis also revealed a significant main effect of language (F(2,46)=12.82, p<.001). Descriptively, RTs were shorter when the background was produced in Irish (1,219 ms) and French (1,243 ms), and they were significantly longer with Italian (1,378 ms). The interaction between these 2 factors was tendential (F(2,46)=2.68, p=.08). Post-hoc comparisons with the HSD Tukey test indicated that with babble, a significant difference emerged between Irish and Italian (p<.001) and between French and Italian (p=.02). Finally, the effect of background was present for Italian (p<.005) and French (p<.05), with RTs significantly faster with fluctuating noise than with babble.

As for Experiment 1, we observed the expected effect of background. Moreover, for French and Italian, i.e. the known languages to the participants, a significant difference between the 2 kinds of backgrounds (fluctuating noise vs. babble) emerged, revealing that acoustic and linguistic information contained in the Italian and French babbles both disrupted the lexical access of the French target words. These results also showed that despite interferences of same nature, the Italian and French babbles did not lead to equal RTs; they were significantly longer with babble spoken in Italian rather than in French.

2.2.3. Comparison of the 2 experiments

We compared the results obtained by the 2 groups of participants in an analysis of variance (ANOVA) that considered the factor group (participants from Exp. 1 vs. participants from Exp. 2) as an independent variable, RT as the dependent variable, and background (babble vs. fluctuating noise) and language (Irish vs. Italian vs. French) as within subjects factors.

The main effect of background emerged significant (F(1,52)=27.67, p<.001), as well as the main effect of language (F(2,104)=16.37, p<.001). However, the main effect of group was not significant (F(1,52)=1.8, p=.18). The interaction between the factors group and language was significant (F(2,104)=3.23, p<.05). The interactions between the factors group and background and the interaction between the factors background and language were tendential, respectively (F(1,52)=3.11, p=.08) and (F(2,104)=2.62, p=.08). Finally, the interaction between all factors was tendential (F(2,104)=2.66, p=.07). Post-hoc comparisons with the HSD Tukey test indicated that mean RTs did not differ between the 2 groups of participants, except with the Italian babble (p=.0002). In this listening condition, RTs were significantly longer for the participants speaking Italian as L2. This significant difference is due to the fact that for the participants from Experiment 1, Italian was an unknown language so only acoustic information disrupted the lexical access of the French target words, whereas for the participants
of the Experiment 2, Italian was intelligible thereby acoustic and linguistic information from the Italian babble both entered in competition with the target words.

3. Discussion and Conclusion

This study examined the speech-in-speech comprehension by using an online measure, i.e. a lexical decision task, to further explore the nature of the interferences which occur between the background speech and the target speech. In order to dissociate the effects of the linguistic and acoustic information from the background speech, we generated fluctuating noise that produces similar acoustic interference as the 4-talker babble but which do not lead to linguistic interference. The significant main effect of background observed in the 2 experiments confirmed that the lexical access of the target words was more disrupted when the background was composed of babble, i.e. with acoustic and linguistic information together, than with background containing fluctuating noise, i.e. with acoustic information alone.

The comparison of the RTs obtained with the 2 kinds of backgrounds in each language manipulated (Irish, Italian and French), highlighted the fact that with babble produced in a known language to the participants (French in Experiment 1, French and Italian in Experiment 2), acoustic and linguistic information both entered in competition during target word identification. With babble spoken in an unknown language (Italian and Irish in Experiment 1, Irish in Experiment 2), acoustic information alone disrupted the lexical access of the target words. Thus, with a known language, the linguistic content of the babble is intelligible and competes with linguistic information of the target speech, leading to longer RTs.

To date, studies that have manipulated the language of the background speech have reported that a babble produced in an unknown language to the participants always leads to better performances than a babble spoken in a known language [3,4,5]. Data of Experiment 1 which explored the nature of the interferences of 2 unknown languages (Irish and Italian), revealed that the speed of lexical access varied depending on the unknown language spoken in the background. In fact, expected results were obtained only for Irish, as its RTs were significantly faster than those of French (i.e. the native language). Italian, an unknown language to the participants, led to RTs as long as French even if their interferences were of different nature, i.e. acoustic interference for Italian, linguistic and acoustic interferences for French. We also observed that the linguistic distance of Irish and Italian with French did not influence the performances as no linguistic interference was observed for Irish and Italian.

In the second experiment, we further explored the effect observed between the Italian and French babbles in the Experiment 1 by testing the same experimental conditions with French native participants speaking Italian as L2. Results clearly showed that the fact that Italian was an intelligible language to the participants led to linguistic interference from Italian babble. Indeed, RTs were significantly longer for the participants speaking Italian as L2 (Experiment 2) than for the monolingual participants (Experiment 1). Moreover in Experiment 2, we can also note that RTs were significantly longer with the Italian babble than with the French babble. In other words, the listening situation was more difficult when the French target words were masked by background speech produced in their L2 (Italian) than in their native language (French). An explanation would be that listening to linguistic information spoken in a native language (target words) and in a L2 (babble) activates the lexicon of each of these 2 languages. Consequently, the effect of linguistic interference would be larger in this listening condition than when the target speech and the concurrent speech are produce in the native language. This would be in line with a work that explored bilingual language spoken processing both in auditory modality and in visual modality by using the eye-tracking technique [11]. Russian native participants speaking English as L2 heard target objects names at the end of an instruction spoken in English or in Russian, and then they were asked to identify these target objects from a set of items in a visual display. Results indicated more eye movements for the objects whose name was in the opposed language from the target language than in the same language (English vs. Russian names), suggesting strong competition between the two languages. Also it appears, as in our results, that the effect of interference inter-languages is stronger than the intra-language one. Further experiments are needed to confirm this assumption.

4. Acknowledgement

This research was supported by a European Research Council grant to the SpiN project N°209234, and by the French “Direction Générale de l’Armement”.

5. References


