Multi-talker background and semantic priming effect

Marie Dekerle, Véronique Boulenger, Michel Hoen, Fanny Meunier

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 31 October 2014
doi: 10.3389/fnhum.2014.00878

Multi-talker background and semantic priming effect

***Marie Dekerle*[1,2]*, *Véronique Boulenger*[2,3], *Michel Hoen*[2,4] *and Fanny Meunier*[1,2]**

[1] *Laboratoire sur le Langage, le Cerveau et la Cognition, Centre National de la Recherche Scientifique, UMR 5304, Lyon, France*
[2] *University of Lyon, Lyon, France*
[3] *Laboratoire Dynamique Du Langage, Centre National de la Recherche Scientifique, UMR5596, Lyon, France*
[4] *Centre National de la Recherche Scientifique, UMR5292, Institut National de la Santé et de la Recherche Médicale, U1028, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, Lyon, France*

Edited by:
Sonja A. E. Kotz, Max Planck Institute Leipzig, Germany

Reviewed by:
Carolyn McGettigan, Royal Holloway University of London, UK
Frederic Marmel, Universidad de Salamanca, Spain

***Correspondence:**
Marie Dekerle, Groupe de Recherche ALP, Laboratoire sur le Langage, le Cerveau et la Cognition, Institut des Sciences Cognitives, 67, Bd Pinel, 69675 BRON CEDEX, France
e-mail: marie.dekerle@isc.cnrs.fr

The reported studies have aimed to investigate whether informational masking in a multi-talker background relies on semantic interference between the background and target using an adapted semantic priming paradigm. In 3 experiments, participants were required to perform a lexical decision task on a target item embedded in backgrounds composed of 1–4 voices. These voices were Semantically Consistent (SC) voices (i.e., pronouncing words sharing semantic features with the target) or Semantically Inconsistent (SI) voices (i.e., pronouncing words semantically unrelated to each other and to the target). In the first experiment, backgrounds consisted of 1 or 2 SC voices. One and 2 SI voices were added in Experiments 2 and 3, respectively. The results showed a semantic priming effect only in the conditions where the number of SC voices was greater than the number of SI voices, suggesting that semantic priming depended on prime intelligibility and strategic processes. However, even if backgrounds were composed of 3 or 4 voices, reducing intelligibility, participants were able to recognize words from these backgrounds, although no semantic priming effect on the targets was observed. Overall this finding suggests that informational masking can occur at a semantic level if intelligibility is sufficient. Based on the Effortfulness Hypothesis, we also suggest that when there is an increased difficulty in extracting target signals (caused by a relatively high number of voices in the background), more cognitive resources were allocated to formal processes (i.e., acoustic and phonological), leading to a decrease in available resources for deeper semantic processing of background words, therefore preventing semantic priming from occurring.

Keywords: semantic priming, informational masking, cocktail party, cognitive load, effortfulness hypothesis

INTRODUCTION

In daily life, speech is rarely perceived in silence, but with interference from wind, music or other people's conversation. Although often used to study psychoacoustic topics (Brungart, 2001; Brungart et al., 2001; McDermott, 2009), speech-in-noise and cocktail party situations (i.e., speech-in-speech, Cherry, 1953) also appear to be interesting paradigms to tackle linguistic processes and competition occurring between backgrounds and targets (Hoen et al., 2007; Boulenger et al., 2010). Our study aimed to investigate the extent to which multi-talker background is processed semantically when listening to speech-in-speech and therefore how the cocktail party situation can be used to study the automaticity of word semantic activation.

The cocktail party situation is described as involving two types of masking effects: energetic and informational masking (Brungart, 2001). Energetic masking relies on the spectro-temporal features of sounds and results from different sounds stimulating the same part of the cochlea at the same time so that one of them cannot be heard (i.e., as two signals increasingly share spectro-temporal characteristics, energetic masking becomes more efficient). In multi-talker background situations, the magnitude of energetic masking is proportional to the number of voices that comprise the background (Simpson and Cooke, 2005). Informational masking, however, usually refers to masking effects that cannot be attributed to energetic masking. Specifically, it is related to the overlap of information carried by the different signals at a higher level (e.g., lexical level and working memory; see Durlach, 2006; Cooke et al., 2008; Mattys et al., 2009; Mattys and Wiget, 2011). Whereas background noise mainly elicits energetic masking, a speech background produces both energetic and informational masking (Brungart et al., 2006). Despite the masking, it is still possible to detect and recognize a word or linguistic token embedded in a babble. Of course, as more voices are present in the babble, participants become less accurate (Freyman et al., 2004). However, it is interesting to note that Simpson and Cooke (2005) showed, using a—6 dB SNR, that intelligibility decreases as a monotonic function of the number of speakers in babbles of up to 8 voices. Specifically, participants' accuracy to detect the target token decreases as the number of voices increases up to 8 voices. Further increasing the number of voices does not lead to a decrease in accuracy. These results suggest that if energetic masking is too high, informational masking decreases with the

diminution of the available linguistic cues. For example, with more than 8 talkers, phonetic cues are not or less available and therefore cannot be attributed incorrectly to the target.

The first aim of this paper is to test whether semantic features are involved in informational masking. It has been established that informational masking is not monolithic and occurs at many linguistic levels. Indeed, a multi-talker background will create less interference on a target word, if it is pronounced in a different language (Van Engen and Bradlow, 2007; Brouwer et al., 2012) and different languages will not have the same masking power (Gautreau et al., 2013). By manipulating the number of talkers in the background, Boulenger et al. (2010) revealed lexical competitions between a 2-talker background and target speech using a lexical decision task. Increasing the number of voices in the background, however, led to the disappearance of lexical interference because of increased energetic masking (i.e., words from the background became less intelligible and therefore competed less with target processing). However, using the same paradigm but with an intelligibility task, Hoen et al. (2007) showed that lexical processing of a background could be performed with up to 4 concurrent voices; beyond that threshold, masking was too high and seemed to prevent linguistic processes. Although it has been shown that phonological and lexical information contribute to the informational masking effect, our experiments tested the role of semantic information.

Processing of the background's semantic content has already been highlighted with 2 talkers, pronouncing either semantically correct sentences (i.e., "rice is often served in round bowls") or incorrect sentences (i.e., "the great car met the milk," Brouwer et al., 2012). Semantic incoherence in the background impacts the recognition of the target sentence. This result suggests that the background signal with 2 talkers is processed semantically. Our experiments aimed to identify how many talkers are allowed in this semantic processing using words and how semantic information from the backgrounds interferes with the identification of target words.

The ability to semantically process auditory words presented outside of the attentional focus is traditionally studied using dichotic listening. This paradigm allows to study pure informational masking as no energetic masking occurs in dichotic listening. However, discrepant results have been reported (Cherry, 1953; Lewis, 1970; Eich, 1984; Wood et al., 1997; Dupoux et al., 2003). In 1984, Eich showed a semantic effect on the recognition of words presented in the unattended channel. However, this effect resulted, at least partially, from an attentional shift toward the to-be-ignored channel (as suggested by Wood et al., 1997). As the speaker rate was very slow in Eich's experiment (85 words per minute), it allowed participants to listen to the supposedly unattended channel without disturbing the primary task (in the case of Eich's study, a shadowing task). In replicating Eich's study using the same speech rate, Wood et al. (1997) observed the same semantic effect; however, it disappeared if the speaker's rate was increased to 170 words per minute, corresponding to a more ecologically valid rate. The authors concluded that as this faster speech rate demanded more cognitive resources, participants could no longer shift attention to the unattended channel while performing the primary task, suggesting that at least in dichotic listening, informational masking does not involve semantic information. The issue raised by this paradigm is that the spatial separation of auditory signals creates a masking release compared to a binaural condition and therefore facilitates stream segregation that could prevent competition between the to-be-ignored and target speech (Drullman and Bronkhorst, 2000; Hawley et al., 2004).

Concerning semantic activation and according to traditional theoretical models, semantic memory is organized into networks. The recognition of one word leads to its activation in semantic memory, and this activation is supposed to spread automatically to other related concepts (Collins and Quillian, 1969; Collins and Loftus, 1975). This supposition is derived from semantic priming paradigms, shown in auditory and visual modalities, in which the presentation of prime word leads to faster recognition of a semantically related target word (Meyer and Schvaneveldt, 1971; Donnenwerth-Nolan et al., 1981; Radeau, 1983; Schacter and Church, 1992; Radeau et al., 1998; Spruyt et al., 2012). For example, the presentation of the prime "nurse" before the target "doctor" facilitates the recognition of the target word "doctor" compared to a condition in which the prime is unrelated to the target (Meyer and Schvaneveldt, 1971). Adapting this paradigm to the cocktail party situation will allow us to investigate if the semantic content of the background is processed and interferes with the target word, despite decreased intelligibility. Some background words will therefore act as primes.

In the current study, we used the rationale of a priming paradigm by manipulating the association between words pronounced in the background and target words. Additionally, we varied the amount of masking to evaluate how it modulates semantic priming effects. Participants were required to perform a lexical decision task on a target item (i.e., decide whether the target item is a word or a pseudo-word) embedded in backgrounds composed of 1 to 4 voices depending on the experiment. These voices could pronounce words that were semantically related to each other and that were related or unrelated to the target. They acted as primes and were called Semantically Consistent (SC) voices. Additional voices pronounced words that were always unrelated to each other and unrelated to the target, acting as maskers. They were called Semantically Inconsistent (SI) voices.

Across experiments, we manipulated the ratio between SC and SI voices. The aim was to test the preservation of the semantic processing of SC voices despite increased masking (i.e., more SI voices). In Experiment 1, backgrounds were composed of 1 or 2 SC voices. In Experiments 2 and 3, respectively, 1 and 2 SI voices were added to each background to increase masking and therefore, decrease the intelligibility of the SC voices. Consequently, in Experiment 2, backgrounds in one condition consisted of 1 SC voice and 1 SI voice and in a second condition of 2 SC voices and 1 SI voice. In Experiment 3 they comprised 1 SC voice and 2 SI voices in one condition and 2 SC voices and 2 SI voices in the other condition.

Overall increasing the number of voices allowed us to examine if and how semantic priming can be impacted by the increase in the number of talkers in the background. Additionally, the variation in the number of SC voices compared to the number of SI voices allowed us to study the effect of prime saliency on semantic

processing and therefore its participation in informational masking. Indeed, across experiments, backgrounds can consist of the same number of voices whereas the number of SC voices compared to the number of SI voices could differ (e.g., 3 voices in the background: either 2SC/1SI in Experiment 2 or 1SC/2SI in Experiment 3).

If semantic processing can occur automatically, semantic priming should be observed at least as long as background words are intelligible and should not be disturbed by increased masking and decreased prime saliency. Indeed, automaticity is defined as a strategy free processing that occurs without using the resources of a limited capacity central processor (Neely, 1977). Therefore, if semantic processing is strategy free, it should occur even if participants are not aware that a given word is presented to them (as is done in visual modality in classical masked priming paradigms, see Forster and Davis, 1984).

EXPERIMENT 1

The aim of this experiment was to first establish set up and test our paradigm and experimental materials. Backgrounds were composed of 1 or 2 SC voices that pronounced words sharing semantic features with each other. In the related condition, target words belonged to the same semantic field as the prime, but they did not in the unrelated condition. We therefore expected to observe a semantic priming effect: participants should more quickly and accurately identify target words in the related compared to the unrelated condition. The second aim of this first experiment was to test if the presence of 2 voices in the background would affect participants' performance as suggested by the psychoacoustic literature (Brungart, 2001; Brungart et al., 2001). We therefore hypothesized that target words would be answered to more slowly and less accurately in the 2SC condition compared to the 1SC condition. Finally, we examined whether the semantic priming effect was modulated by increased energetic and informational masking caused by the augmentation of the number of voices in the background.

METHOD

Participants

Twenty-seven participants (20 females) volunteered for this experiment. All were right-handed, French native speakers and reported no known hearing or language disorder. Subjects' ages ranged from 18 to 25 years old. All participants gave written informed consent and were not aware of the experiment's purpose. They were compensated for their participation. The protocol that was used in this experiment was approved by the local ethics committee (CPP Sud-Est IV, Lyon; ID RCB: 2008-A0 0708-47).

Stimuli

Forty-eight disyllabic target words ($M_{\text{lexical frequency}} = 21.94$ per million, $SD = 18.75$ according to the French database Lexique 3, New et al., 2001) were selected, and each word belonged to a specific semantic field (e.g., *CAROTTE* "carrot"; *MÉTRO* "subway"). Each target word was matched to 10 words belonging to the same semantic field (e.g., *CAROTTE* "carrot" was associated with *légume, chou, céleri, salade, tomate* "vegetable, cabbage, celery, lettuce, tomato"). As participants had to perform a lexical decision task, 48 pseudo-words respecting French phonotactic rules were created (e.g., *PLARO*, *HUMEL*). Ten words sharing semantic features with each other were arbitrarily associated with each pseudo-word target, resulting in a total of 96 groups of 10 words (See Supplementary Material) ($M_{\text{lexical frequency}} = 21.86$, $SD = 18.20$). As each background comprised 1 or 2 SC voices (related or not to the target), each group was divided into two subgroups of 5 words one of the subgroups was spoken by a first speaker (S1), and the other by a second speaker (S2).

Target words were presented with a semantically related (related condition) or semantically unrelated background (unrelated condition). In the unrelated condition, SC voices pronounced words that were semantically related to each other but not to the target (see **Figure 1**). Backgrounds comprised 1 SC voice (1SC condition) or 2 SC voices (2SC condition). The 48 target words were divided into 4 groups of 12 words, the mean frequency did not differ significantly between the groups ($F < 1$), nor did the number of phonemes [$M = 6.97$, $SD = 5.65$; $F_{(3,\ 44)} = 1.1$, *n.s.*] and phonological neighbors [$M = 4.75$, $SD = 0.81$; $F_{(3,\ 44)} = 2.2$, *n.s.*]. Each group of twelve target words was assigned to a condition (1SC related, 1SC unrelated, 2SC related, 2SC unrelated) depending on the experimental list. The same was true for pseudo-words. Four experimental lists of 96 stimuli (i.e., 48 target words and 48 target pseudo-words) were created so that each target word was presented in each condition, but only once in a list (each participant was presented with one list only).

Targets and SC voices were recorded by 3 different French native female speakers (age: 21–22) in a sound-proof room (22 kHz, mono, 16 bits). Auditory sequences of 5 words from Speakers 1 and 2 (S1 and S2) were segmented into 3 s periods. The periods were then normalized at an intensity of 60 dB-A and mixed together to create backgrounds. All audio files were synchronized at the beginning, so all voices started to speak at the same time. However, as all voices pronounced words of different lengths, they soon became desynchronized, and there was always one speaker talking in the background. Targets recorded by Target Speaker (TS; also normalized at an intensity of 60 dB-A) were inserted 2 s after the start of the backgrounds (so that each participant always had the same exposure to the background before the target speech was presented), with a 0 dB SNR (Signal/Noise Ratio; see **Figure 1**). Because the backgrounds, which comprised 1 or 2 voices, generated different amounts of energy, the intensity of all stimuli was varied over a ± 3 dB range in 1 dB steps to prevent participants from predicting condition depending on individual stimuli intensity.

Procedure

Participants sat in front of a computer screen and heard the stimuli binaurally through headphones at a comfortable level (mean level 65 dB-A, ranging from 62 dB-A to 68 dB-A, normalized using an artificial ear). A fixation cross was presented on the screen at the beginning of each trial and remained on the screen during stimulus presentation. Participants were asked to listen to the stimuli to decide as quickly and accurately as possible whether the target was a word or a pseudo-word, by pressing one of two

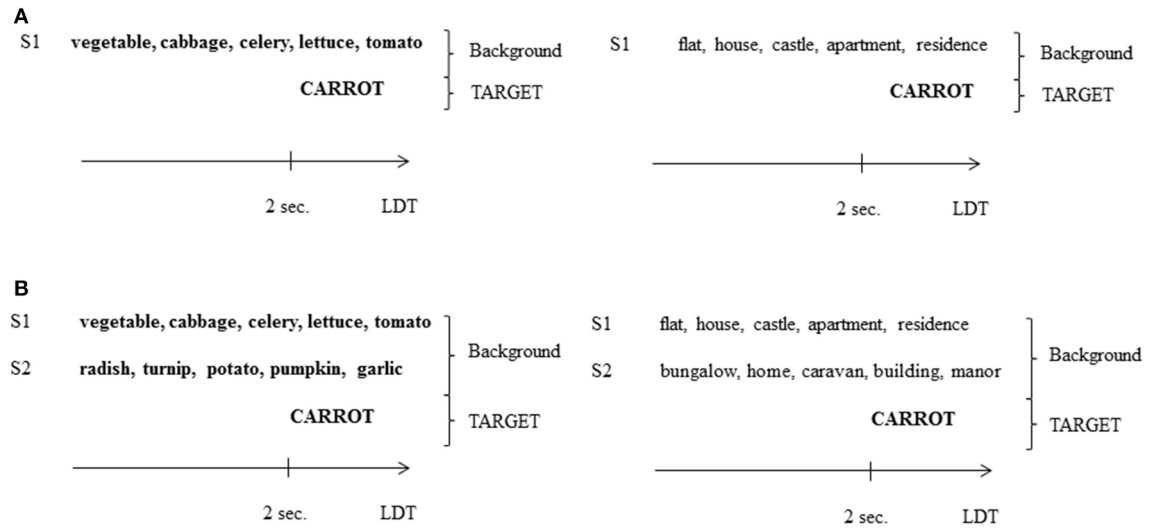

FIGURE 1 | (A) Example of two backgrounds in the 1SC condition of Experiment 1, presented with a semantically related target word (left; related condition) or not (right; unrelated condition). S1, speaker 1, LDT, Lexical Decision Task. **(B)** Example of a background in the 2SC condition of Experiment 1, presented with a semantically related target word (left; related condition) or not (right; unrelated condition). S2, speaker 2.

pre-specified keys. After a response was given, a string of hash marks indicated that the trial was over; participants could then press a key to start the next trial. Half of the participants gave the response to "word" with their left hand and to "pseudo-word" with their right hand. As all participants were right-handed, they might answer faster with their right hand than their left hand. To avoid this confounding effect, the other half were given the opposite instruction. A training session composed of twelve trials (different from the experimental stimuli) preceded the test session so that participants could acclimate to the stimuli and the task.

RESULTS

Two Two-Way repeated measures analyses of variance (ANOVAs) by participants (F_1) and by items (F_2) were conducted, with Response Times (RTs, in ms) and Error Rates (ERs) for target word identification as dependent variables. We included Number of Voices in the background (1 Voice, 1SC or 2 Voices, 2SC) and Semantic Link between prime and target (related or unrelated) as within-subjects factors. Three participants were excluded from analyses because of very high ERs (more than 40%). Four target words error rates greater than 50% were also excluded from Item analyses (*POIGNET*, *RIDEAU*, *RATON*, and *RACINE* "wrist, curtain, baby rat, root"). Trials with RTs below or above 2.5 standard deviations from the individual means (4.5%) and trials in which participants made mistakes (19.5%) were not included in RTs analysis. Means and Standard Deviations (SDs) of RTs and ERs are summarized in **Table 1**.

The ANOVA by participants first revealed a significant main effect of the Number of Voices: participants were faster [$F_{1(1,\ 23)} = 4.25$, $p = 0.05$] and more accurate [$F_{1(1,\ 23)} = 9.12$, $p < 0.01$] to identify targets in the 1SC condition ($M_{RT} = 1008$ ms, $SD = 157$; $M_{ER} = 11.7\%$, $SD = 10.9$) than in the 2SC condition ($M_{RT} = 1042$ ms, $SD = 161$; $M_{ER} = 20.1\%$, $SD = 16.5$). The Item analysis, however, did not highlight an effect of the Number of Voices on RT [$F_{2(1,\ 43)} = 1.9$, $p = 0.1$], although target words were better categorized as words in the 1SC condition [$F_{2(1,\ 43)} = 13.43$, $p < 0.001$].

Table 1 | Means and Standard Deviations (SDs) of Response Times (RTs) and Error Rates (ERs) depending on the number of voices in the background and the semantic link between prime and target in Experiment 1.

		1SC		2SC	
		Related	**Unrelated**	**Related**	**Unrelated**
RT (ms)	Mean	980	1035	1014	1070
	SD	170	142	169	150
ER (%)	Mean	7.08	16.33	19.5	20.77
	SD	6.82	12.34	19.04	14.05

1SC, 1 SC voice condition; 2SC, 2 SC voices condition; related, semantic link between the prime and the target; unrelated, no semantic link between the prime and the target.

The main effect of Semantic Link also appeared to be significant on RTs [$F_{1(1,\ 23)} = 14.24$, $p < 0.001$; $F_{2(1,\ 43)} = 4.63$, $p < 0.05$], participants responded faster if targets shared semantic features with the prime ($M = 997$ ms, $SD = 169$) than if they did not ($M = 1053$ ms, $SD = 145$); this resulted in a 56 ms priming effect. This effect was also significant for ERs in the participant analysis [$F_{1(1,\ 23)} = 3.93$, $p = 0.05$], and there was only a trend in the item analysis [$F_{2(1,\ 43)} = 3.07$, $p < 0.10$]. Participants tended to be more accurate in the related condition ($M = 15.4\%$, $SD = 15.4$) than in the unrelated condition ($M = 18.55\%$, $SD = 13.2$). There was no significant interaction between the two factors for RTs ($F_1 < 1$; $F_2 < 1$) and ERs [$F_{1(1,\ 23)} = 2.34$, *n.s.*; $F_{2(1,\ 43)} = 1.97$, *n.s.*], suggesting that the semantic priming effect was not

modulated by the Number of Voices (one or two) in the background.

DISCUSSION

These results first highlight that participants were slowed by the increase in the number of voices in the background. This effect is certainly attributable to enhanced target masking in the two-voice condition (Brungart, 2001; Brungart et al., 2001). Interestingly, participants' performance was improved by the semantic relationship between the prime and target, and this effect was independent of the number of voices, suggesting that the increase in masking from one to two background voices, was not sufficient to prevent semantic processing. However, in this first experiment, prime was salient in both conditions (1SC voice and no SI voice or 2SC voices and no SI voice). To test whether participants could still take advantage of the semantic relationship between target and prime if the intelligibility of the SC voices was further decreased, we conducted a second experiment in which a SI voice was added to each background.

EXPERIMENT 2

This second experiment aimed to investigate whether the semantic priming effect would resist increased masking. An SI voice was therefore added to each background. This voice pronounced words sharing no semantic features with each other or with the target word, whatever the condition. The purpose was to use the same material and procedure as in Experiment 1 with the addition of mask on the SC voices. In Experiment 2, backgrounds were composed of two voices (1 SC voice + 1 SI voice) or 3 voices (2 SC voices + 1 SI voice). A deleterious effect of the number of voices on participants' performance was predicted, and we expected that the presence of SI voice would not affect the semantic priming effect if this latter effect is automatic.

Another change was made in Experiment 2 regarding target items. In Experiment 1, target items were pronounced by a female speaker, and were consequently, difficult to detect among the other female speakers (S1 and S2). These difficulties might partly explain the low accuracy and long response times to target words inserted in babbles that were only composed of one or two voices. To avoid flux segregation difficulties (Festen and Plomp, 1990; Brungart et al., 2001), target items were therefore pronounced by a male speaker (Target Speaker 2; TS2) in the two following experiments.

METHOD

Participants

Twenty-four right-handed French native speakers (18 females), aged 18–30 years, participated in this second experiment. They had no known auditory or language disorders. All participants gave their written informed consent and were compensated for their participation. None of the participants had been tested in Experiment 1 and they were not aware of the aim of the study before testing.

Stimuli

To add an SI voice to each background used in Experiment 1, 96 groups of 5 words ($M_{\text{lexical frequency}} = 18.15$, $SD = 9.75$), not semantically related to each other, were generated. Average lexical frequency did not differ between SC voices (from Experiment 1) and the SI voice as shown by an ANOVA [$F_{(2,\ 190)} = 1.16$, *n.s.*]. Each group was selected to mask a specific prime (composed of 1 or 2 SC voices), which shared no semantic link (e.g., the prime *légume, chou, céleri, salade, tomate* "vegetable, cabbage, celery, lettuce, tomato" was always presented with the SI voice pronouncing *policier, intéressant, cour, affiche, étagère* "policeman, interesting, yard, poster, shelf"). This SI voice was recorded by another French native female speaker (S3, age = 20) using the same method as in Experiment 1.

Backgrounds composed of 1 SC voice (S1) in Experiment 1 were now composed of 1 SC voice and 1 SI voice (S1 + S3), corresponding to the 1SC/1SI condition, and backgrounds composed of 2 SC voices (S1 + S2) in Experiment 1 were now composed of 2 SC voices and 1 SI voice (S1 + S2 + S3), corresponding to the 2SC/1SI condition. The 4 groups of 12 target words and pseudo-words created for Experiment 1 were used. Target words were presented in each of the 4 conditions: 1SC/1SI related, 1SC/1SI unrelated, 2SC/1SI related and 2SC/1SI unrelated. The corresponding number of pseudo-words was also presented. Four experimental lists were created so that each target word was seen in each condition but only once in a list.

Recordings of S1 and S2, used in the previous experiment, were mixed with S3 following the previously established experimental lists. Targets were recorded by a French native male speaker (Target Speaker 2; age = 20) and were inserted into backgrounds 2 s after the start of the sequence (see **Figure 2**).

Procedure

The procedure was the same as in Experiment 1.

RESULTS

Similar analyses as in Experiment 1 were performed by subjects (F_1) and by items (F_2), with RTs (in ms) and ERs for target word identification as the dependent variables. We included Number of Voices in the background (two voices, 1SC/1SI or three voices, 2SC/1SI) and Semantic Link between prime and target (related or unrelated) as within-subjects factors. As in Experiment 1, target words for which less than 50% of participants answered correctly were not analyzed. The target words *BILLET* and *CHIGNON* ("ticket" and "bun") were therefore not included. Moreover, 15% of data were excluded (13% of errors and 2% of extremes values) from RT analysis. Mean RTs and ERs with SDs are summarized in **Table 2**.

The Two-Way repeated measures ANOVAs showed a significant main effect of the Number of Voices in the backgrounds [$F_{1(1,\ 23)} = 6.08$, $p < 0.05$; $F_{2(1,\ 45)} = 4.34$, $p < 0.05$]. Participants responded faster to target words if backgrounds comprised 2 voices ($M = 946$ ms, $SD = 157$) compared to 3 voices ($M = 973$ ms, $SD = 175$). This main effect was also significant for ERs [$F_{1(1,\ 23)} = 7.18$, $p < 0.05$] but only in the participants' analysis [$F_{2(1,\ 45)} = 3.72$, $p < 0.10$]. Responses were more accurate in the 1SC/1SI condition ($M = 8.6\%$, $SD = 8.2$) than in the 2SC/1SI condition ($M = 12.5\%$, $SD = 9.1$).

The main effect of Semantic Link was also significant [$F_{1(1,\ 23)} = 5.32$, $p < 0.05$; $F_{2(1,\ 45)} = 5.38$, $p < 0.05$], responses

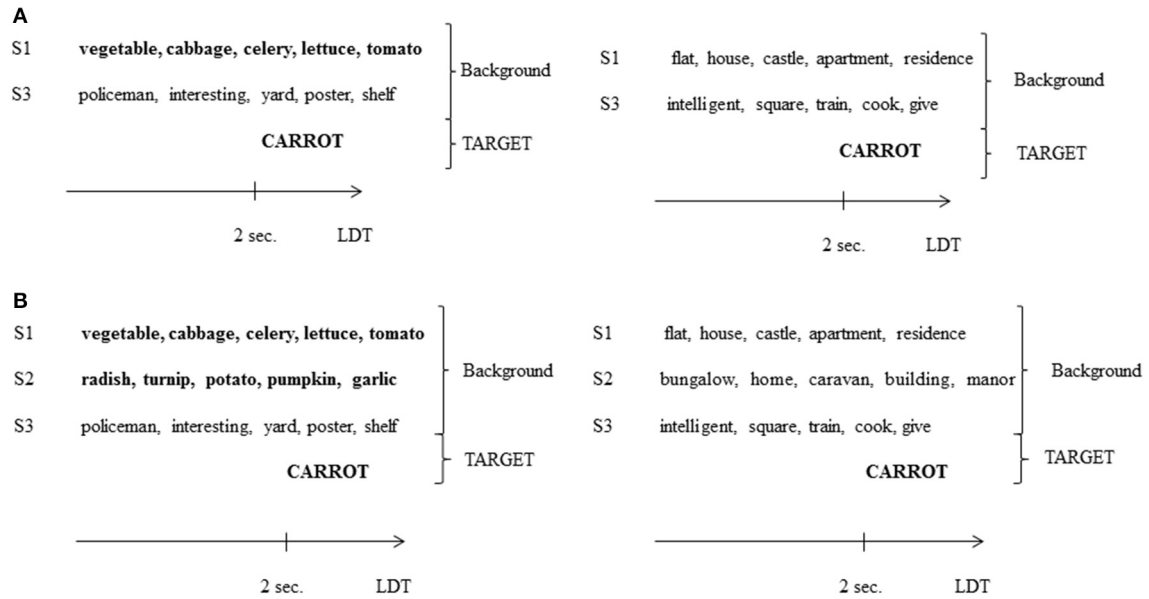

FIGURE 2 | (A) Example of a background in the 1SC/1SI condition of Experiment 2, presented with a semantically related target word (left; related condition) or not (right; unrelated condition). S3, speaker 3 (see the legend of **Figure 1** for other abbreviations). **(B)** Example of a background in the 2SC/1SI condition of Experiment 2, presented with a semantically related target word (left; related condition) or not (right; unrelated condition).

Table 2 | Means and SDs of RTs and ERs depending of the number of voices in the background and the semantic link between prime and target in Experiment 2.

		1SC/1SI		2SC/1SI	
		Related	**Unrelated**	**Related**	**Unrelated**
RT (ms)	Mean	933	958	945	1000
	SD	165	153	162	187
ER (%)	Mean	6.47	10.7	10.7	14.33
	SD	7.29	9.14	9.32	8.78

1SC/1SI, 1 SC voice and 1 SI voice condition; 2SC/1SI, 2 SC voices and 1 SI voice condition.

were 40 ms faster if the prime and target shared semantic features ($M = 939$ ms, $SD = 162$) compared to if they did not share features ($M = 979$ ms, $SD = 170$). This effect also reached significance for ERs in the participants' analysis [$F_{1(1,\ 23)} = 5.33$, $p < 0.05$; $F_{2(1,\ 45)} = 1.98$, $p = 0.10$]. ERs decreased by 3.9% if the prime and target were semantically related ($M = 8.61\%$, $SD = 8.5$ in the related condition and $M = 12.53\%$, $SD = 8.8$ in the unrelated condition).

There was no significant interaction between the Number of Voices in the backgrounds and the Semantic Link between prime and target for RTs [$F_{1(1,\ 23)} = 1.92$, *n.s.*; $F_2 < 1$] and ERs ($F_1 < 1$; $F_2 < 1$), indicating that the priming effect was not affected by the increase in the number of voices in the backgrounds.

DISCUSSION

As in Experiment 1, performances improved if the prime and target were semantically related, although participants were slower in condition 2SC/1SI than 1SC/1SI because of increased masking (Bronkhorst, 2000; Brungart, 2001; Brungart et al., 2001). Interestingly, there was again no indication that increased masking reduced the semantic priming effect. To further test this resistance of the semantic priming effect to prime intelligibility loss, we conducted a third experiment in which a second SI voice was added to each background to further decrease prime saliency. However, the intelligibility of the target item may decrease with the addition of a second SI voice in the background. However, as the target item is pronounced by a male voice and embedded in a female voice background, it is still quite salient compared to background voices (Brungart et al., 2001). Indeed, Brungart and collaborators showed that participants more easily recognize a target sentence if it was embedded in a background composed of voices of a different sex than if all the sentences (background and target) were pronounced by speakers of the same sex. Additionally, in our experiment, target items were presented at the same time (2 s after the beginning of the stimulus), and this regular timing helps participants to detect the target, as they know when to listen to it. Consequently, in our paradigm, the masking effect of the SI voice on the target item was quite low compared to its effect on SC voices.

EXPERIMENT 3

This last experiment was the same as Experiment 2, except that a second SI voice was added to each background (i.e., backgrounds comprised 3 or 4 voices). The same method was used as in Experiment 2. The aim was to further increase the masking of SC voices to test the resistance and automaticity of semantic processing. As the number of voices in the backgrounds increased (i.e., up to 4 voices), we wanted to make sure whether participants could still identify words from the background. Therefore, after the

main experiment, we asked participants to perform a recognition test. This post-test was designed to examine whether, in agreement with Hoen et al. (2007), words in the background were still intelligible. It aimed to clarify, in the case of significant semantic priming, if this effect resulted from preserved intelligibility of the prime words by testing lexical access or from automatic processing. In case of preserved intelligibility we cannot prove that the semantic effect results from automatic processing, it may be either automatic or strategic. However, if a priming effect is found without preserved intelligibility, semantic processing is automatic (as shown with masked priming paradigm in visual modality, see Dehaene et al., 1998; Naccache and Dehaene, 2001; Spruyt et al., 2012). Proof of non-intelligibility was consequently needed in the case of a significant semantic priming effect to provide evidence for an automatic process. Otherwise, we would not be able to detect whether automatic components are present in semantic processes. As in our previous experiments, we expected to observe a significant effect of the number of voices in the background; increasing the number of voices has shown to decrease intelligibility for up to 8 voices (Simpson and Cooke, 2005). No interaction between the number of voices and the semantic association between prime and target was expected, at least if the words composing the background were still intelligible.

METHOD

Participants

Twenty-four participants (19 females) were recruited for this experiment (age: 18–34). All were right-handed French native speakers and reported no hearing or speech disorder. They gave written informed consent and were compensated for their participation. None of them had participated in Experiments 1 or 2, and they were unaware of the experiment's purpose prior to testing.

Stimuli

To create an extra SI voice, pronounced by a fourth speaker (S4), 96 groups of 5 words, that were semantically unrelated to each other, were generated ($M_{\text{lexical frequency}} = 20.88$, $SD = 1.22$). The word mean frequency did not differ between the different voices (SC and SI voices; $F < 1$). As in Experiment 2, each group of words was matched to a prime (composed of 1 or 2 SC voices) with which it did not share any semantic features and was systematically presented with (e.g., the prime *légume, chou, céleri, salade, tomate* "vegetable, cabbage, celery, lettuce, tomato" was always presented with the masker *étui, liberté, drôle, global, sympathie* "case, freedom, funny, global, sympathy"). Therefore, with the addition of this second SI voice, backgrounds composed of 2 voices (S1 + S3) from Experiment 2 became 3-voice backgrounds (S1 + S3 + S4; 1 SC voice and 2 SI voices; 1SC/2SI condition) and 3-voice backgrounds (S1 + S2 + S3) from Experiment 2 became 4-voice backgrounds (S1 + S2 + S3 + S4; 2 SC voices and 2 SI voices; 2SC/2SI condition). The 4 groups of 12 target words and pseudo-words created in Experiment 1 were used and presented in the following conditions: 1SC/2SI related, 1SC/2SI unrelated, 2SC/2SI related, and 2SC/2SI unrelated. Four experimental lists were created so that each target word was presented in each condition but only once in a list.

The second SI voice (S4) was recorded by a French native female speaker (age = 23), using the same procedure as in previous experiments. S1, S2, and S3's auditory sequences were mixed with S4's. Targets used in Experiment 2 (recorded by TS2) were embedded in the backgrounds 2 s after their beginning start (see **Figures 3**, **4**).

Recognition post-test

A recognition test was devised to test whether participants could recognize words previously heard during the experiment. Fifty words were presented to participants after the main experiment on a sheet of paper, 20 had been previously presented in the backgrounds, whereas 30 were new words not used as stimuli in the experiment. A list of new words was generated ($M_{\text{lexical frequency}} = 23.15$, $SD = 48.29$), and their lexical frequency did not significantly differ from that of previously heard words ($M_{\text{lexical frequency}} = 18.47$, $SD = 20.61$; $F < 1$).

Procedure

The same procedure was used as in Experiments 1 and 2. At the end of the experiment, a post-test was given to participants, instructing them to decide (i.e., write down) whether they had heard the given word during the experiment. They were asked to do the best they could but not to think about it too much, and to simply answer what they thought was right.

RESULTS

Test

The same statistical analyses as in Experiments 1 and 2 were performed by participants (F_1) and by items (F_2), with RTs (in ms) and ERs for target word identification as dependent variables. We included Number of Voices in the background (3 voices, 1SC/2SI condition or 4 voices, 2SC/2SI condition) and Semantic Link between prime and target (related or unrelated) as within-subjects factors. Target words answered correctly by less than half of participants were excluded from analyses (*CHIGNON, EPAULE,* and *HIBOU,* "bun, shoulder, owl"). In total, 16.8% of data were excluded from RTs analysis because of errors (15%) or extreme values (1.8%). Means and SDs for RTs and ERs are summarized in **Table 3**.

The analysis by participants revealed a significant main effect of the Number of Voices on RTs [$F_{1(1,\ 23)} = 4.01$, $p = 0.05$; $F_2 < 1$]. Participants more quickly identified target words if backgrounds were composed of 3 voices (1SC/2SI condition; $M =$ 925 ms, $SD = 155$) than if they were composed of 4 voices (2SC/2SI condition; $M = 941$ ms, $SD = 155$). The ERs analysis also showed that participants responded significantly more accurately [$F_{1(1,\ 23)} = 6.19$, $p < 0.05$; $F_{2(1,\ 44)} = 5.80$, $p < 0.05$] in the 1SC/2SI condition ($M = 10.02\%$, $SD = 10.10$) than in the 2SC/2SI condition ($M = 13.21\%$, $SD = 8.95$). The main effect of Semantic Link was not significant for either RTs [$F_{1(1,\ 23)} = 1.51$, *n.s.*; $F_{2(1,\ 44)} = 3.21$, *n.s.*] or ERs [$F_{1(1,\ 23)} = 3.19$, *n.s.*; $F_{2(1,\ 44)} = 2.92$, *n.s.*]. No significant interaction emerged between the 2 factors for RTs [$F_{1(1,\ 23)} = 1.39$, *n.s.*; $F_{2(1,\ 44)} = 1.10$, *n.s.*] and ERs ($F_1 < 1$; $F_2 < 1$). These last results indicate that participants performed better in the lexical decision task, both in terms of speed and accuracy, if the backgrounds were composed of 3

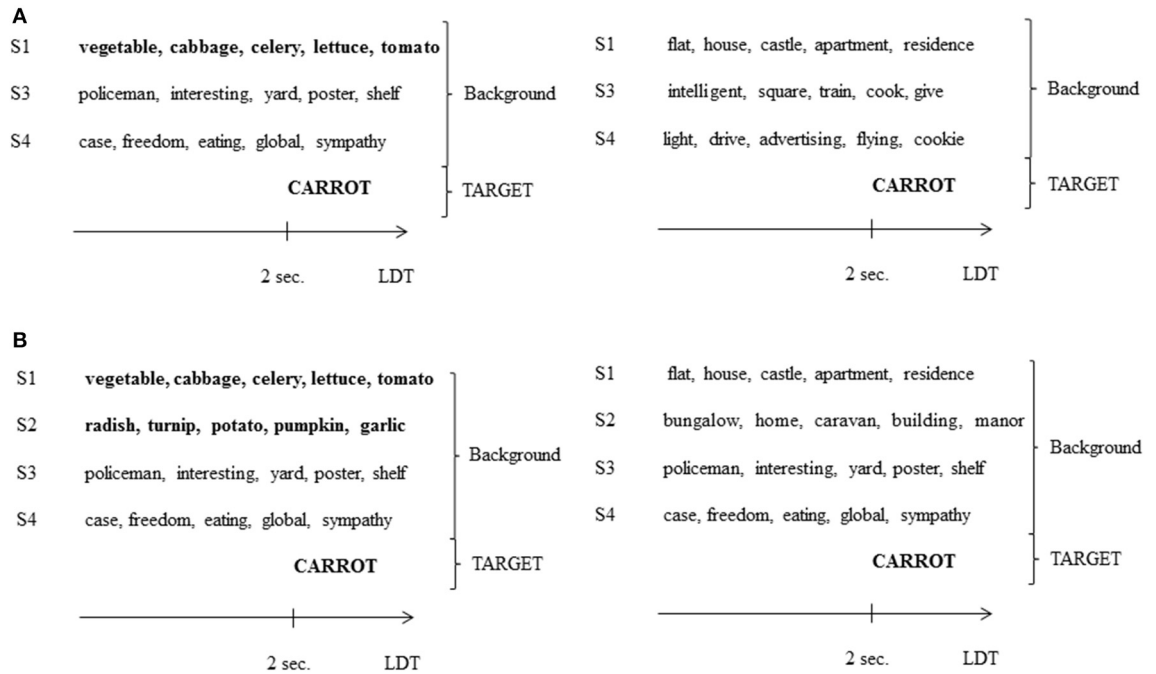

FIGURE 3 | (A) Example of a background in the 1SC/2SI condition of Experiment 3, presented with a semantically related target word (left; related condition) or not (right; unrelated condition). S4, speaker 4 (see the legend of **Figures 1**, **2** for other abbreviations). **(B)** Example of a background in the 2SC/2SI condition of Experiment 3, presented with a semantically related target word (left; related condition) or not (right; unrelated condition).

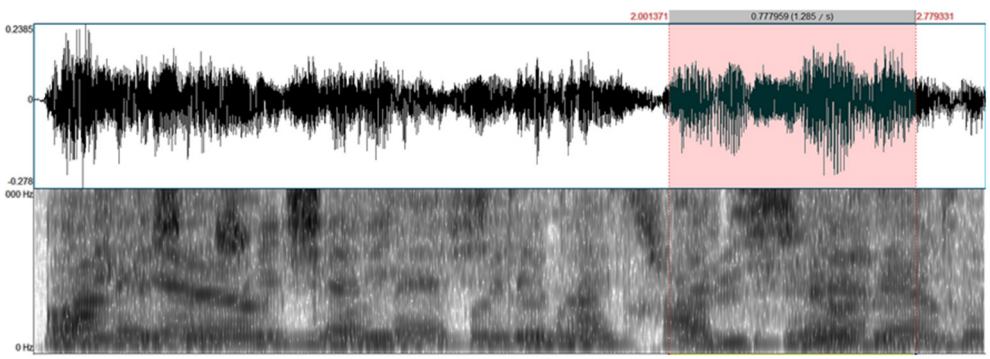

FIGURE 4 | Example of a spectrogram and a wave form of a 2SC/2SI condition stimulus. Red part corresponds to the occurrence of the target word.

voices rather than 4 voices. This finding highlights the impact of masking on target recognition. However, no significant semantic priming effect was observed in these conditions, suggesting that informational masking from 1SC/2SI and 2SC/2SI backgrounds was so efficient that it prevented semantic processing of the prime, which could therefore not affect target word identification. Additionally, in the 2SC/2SI condition, energetic masking was more important; this factor might also explain the important masking of prime and explain the lack of semantic priming.

Post-test

Analysis of the participants' answers showed a mean ER of 44.5% ($SD = 7$) and $d' = 0.26$ ($SD = 0.5$). Although these results are close to chance, both were significant results, as shown by a one-sample t-test (ER $p < 0.01$; d' $p < 0.05$). No significant correlation was found between d' and priming effect ($r = -0.39$, *n.s.*). This finding implies that participants heard some background words, but those words were not used to improve performances. Additionally, a repeated measure ANOVA was performed with ERs as the dependent variable and number of Voices (3 or 4) in the background and Semantic Link between presented word and target as a within subjects factor. Neither the effect of the Number of Voices nor the effect of Semantic Link were significant ($F_{\text{Number of Voices}} < 1$; $F_{\text{Semantic Link}} < 1$).

DISCUSSION

In this third experiment, target word identification was again disturbed by the increase in the number of voices in the

backgrounds, confirming that masking is more efficient in the 4-voice than in the 3-voice condition. Disappearance of the semantic priming effect also suggests that the number of SC voices compared to the number of SI voices was too small. To compare the data of the 3 experiments, we considered the ratio of SC voices over the total number of voices. Therefore, the ratio of SC/total voices is 1 in 1SC and 2SC conditions; ratio 2/3 in 2SC/1SI; ratio 1/2 in 1SC/2SI and 2SC/2SI; and ratio 1/3 in 1SC/2SI. An ANCOVA on all data using the number of voices as the independent variable and the ratio of SC voices/total voices as covariate on priming effect confirmed this hypothesis (effect of ratio: $p < 0.05$): when the ratio was too low, SC voices were not salient enough for participants to perform semantic processing. In a post-test, participants were asked to perform a recognition task immediately after the experiment. Participants scored significantly better than chance at the recognition test, showing that at least some words in the background were identified. This finding is consistent with the results by Hoen et al. (2007) who showed that in a transcription task, with up to 4 voices in the background, participants gave words from the background as responses instead of target words. Overall Experiment 3 showed that participants were unable to process the prime at a semantic level (i.e., no priming effect was observed) although the post-test results suggest that they hear it sufficiently to recognize it in a recognition post-test. This finding suggests that a word can be heard and implicitly encoded without being sufficiently deeply processed to elicit semantic priming.

Table 3 | Means and SDs of RTs and ERs depending on the number of voices in the background and the semantic link between prime and target in Experiment 3.

		1SC/2SI		2SC/2SI	
		Related	Unrelated	Related	Unrelated
RT (ms)	Mean	924	928	928	953
	SD	157	158	156	157
ER (%)	Mean	7.03	11.1	13.00	15.27
	SD	7.30	9.64	11.69	7.88

1SC/2SI, 1 SC voice and 2 SI voices condition; 2SC/2SI, 2 SC voices and 2 SI voices condition.

POST-HOC ANALYSES

To better analyze the impact of the SC/total voices ratio, we conducted *post-hoc* analyses. A HSD Tukey test showed that up to 4 voices in the background if the ratio of SC/total voices was inferior to 1/2, there was no significant semantic priming effect (Experiment 2. 1SC/1SI; Experiment 3. 1SC/2SI and 2SC/2SI). However, if the ratio was superior to 1/2, semantic priming was significant (Experiment 1: 1SC, $p < 0.05$ Cohen's $d = 0.31$; 2SC, $p < 0.05$, Cohen's $d = 0.35$; Experiment 2. 2SC/1SI, $p < 0.05$, Cohen's $d = 0.31$; cf **Figure 5** and **Table 4**).

GENERAL DISCUSSION

SUMMARY

The aim of this study was to investigate to what extent the semantic content of the multi-talker babble could interfere with the processing of targets in a cocktail party situation. In the three reported experiments, participants were required to perform a lexical decision task on a target item embedded in a multi-talker background. The backgrounds were composed of 1 or 2 SC voices pronouncing words that shared semantic features with each other and could be semantically related to the target or not. The ratio of SC voices over the total number of voices in the background was varied across experiments. In Experiment 1, ratios were 1/1 and 2/2 (i.e., only SC voices were presented in the background).

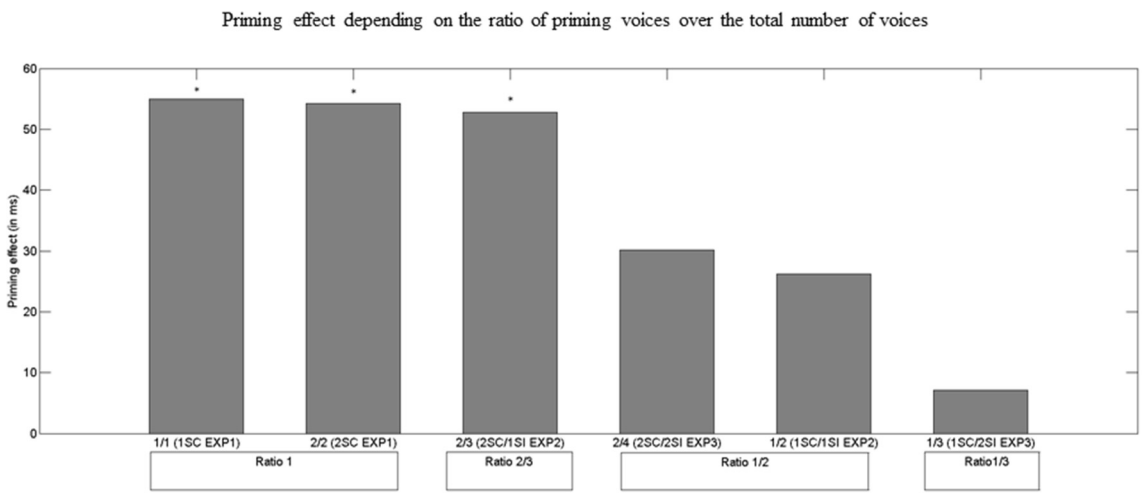

FIGURE 5 | Semantic priming effect (i.e., difference between RTs for unrelated and related targets; in ms) depending on the ratio of SC voices in the background over the total number of voices. 1SC EXP1, 1SC condition in Experiment 1 (1 SC voice; no SI voice); 2SC EXP1, 2SC condition in Experiment 1 (2 SC voices; no SI voice); 2SC/1SI EXP2, 2SC/1SI condition in Experiment 2 (2 SC voices; 1 SI voice); 1SC/1SI EXP2, 1SC/1SI condition in Experiment 2 (1 SC voice; 1 SI voice); and 2SC/2SI EXP3, 2SC/2SI condition in Experiment 3 (2 SC voices; 2 SI voices); 1SC/2SI EXP3, 1SC/2SI condition in Experiment 3 (1 SC voice; 2 SI voices).

Table 4 | Mean priming effect (in ms), SD, and *p*-value for each ratio.

Ratio	1		2/3	1/2		1/3
Condition	1SC	2SC	2SC/1SI	1SC/1SI	2SC/2SI	1SC/2SI
	(EXP1)	(EXP1)	(EXP2)	(EXP2)	(EXP3)	(EXP3)
Mean	55	56	55	25	26	4
SD	70	119	110	88	83	65
P	0.03*	0.03*	0.007*	0.88	0.21	0.99

* *Indicates significant priming. The ratio is expressed in terms of the number of SC voices over the total number of voices in the background: 1, 1 SC voice or 2 SC voices; 2/3, 2 SC voices, 1 SI voice; 1/2, 1 SC voice, 1 SI voice or 2 SC voices, 2 SI voices; 1/3, 1 SC voice, 2 SI voices.*

In Experiments 2 and 3, 1 and 2 SI voices, respectively, that pronounced semantically unrelated words, were added to backgrounds to decrease the intelligibility of the SC voices, which acted as the prime. The ratio of SC/total voices in Experiment 2 was therefore 1/2 and 2/3 and in Experiment 3 it was 1/3 and 2/4.

The main effect of the Number of Voices was significant in each experiment; participants responded faster and more accurately to target words if a smaller number of voices composed the backgrounds. This delayed response time with increasing number of voices is a well-known phenomenon (Brungart et al., 2001; Boulenger et al., 2010). As the number of voices increases, energetic masking is enhanced so that the signal is saturated, and target items are consequently more difficult to process. In addition to physical masking, more information is perceived and must be processed (i.e., informational masking), leading to slower response times.

Overall the main effect of Semantic Link between prime and target seems to be an all-or-none phenomenon. The semantic priming effect could have decreased with an increasing number of voices in the background, but our results suggest this effect does not occur, as no interaction between the Number of Voices and the Semantic Link was observed in any of the three experiments. The main effect of Semantic Link depends on the ratio of SC voices over the total number of voices in the background. A semantic priming effect emerged in our experiments spanning from 1 to 4 voices in the background but only if the number of SC voices was higher than the number of SI voices. This ratio of SC voices is interesting because it highlights the necessity of prime saliency for semantic processing to occur; it also gives an objective measure of this prime saliency across experiments. This finding suggests that informational masking does occur at the semantic level but only if the prime is sufficiently salient. If informational and energetic masking had the same role, a significant semantic priming effect would appear in conditions containing 3 voices (i.e., Experiment 2 2SC/1SI and Experiment 3 1SC/2SI). However, this effect did not occur; we therefore conclude that informational masking can be semantic only if it is sufficiently salient to be used to increase performance.

LEXICAL PROCESSING WITHOUT SEMANTIC ACTIVATION

Our results suggest that semantic processing in cocktail party situation is not automatic. In fact, it seems that in challenging listening situations, one can hear and activate a mental representation of a word without deeply processing it at a semantic level. Background words can be semantically processed with up to 3 voices (2SC/1SI, Experiment 2); however no priming effect emerged in Experiment 3, despite the recognition post-test's results suggesting that primes were heard and recognized. This finding is consistent with the way the language system has been modeled; most models of word processing, both in the auditory and the visual modalities, suggest a distinction between lexical and semantic levels of processing (Marslen-Wilson and Welsh, 1978; McClelland and Elman, 1986; Grainger and Holcomb, 2009). If one considers that these processes are independent stages of word recognition, it seems reasonable that one of these stages can be reached (i.e., lexical) without strongly activating the deepest stage (i.e., semantic).

Many studies have highlighted the automaticity of semantic processing using masked semantic priming (Naccache and Dehaene, 2001; Klauer et al., 2007; Kouider and Dehaene, 2009; Spruyt et al., 2012). However, masked semantic priming is only found in very specific conditions in the visual modality and with semantic categorization tasks (see Van den Bussche et al., 2009, for a review). One can therefore assume that semantic activation remains superficial in masked priming paradigm, and only activates very close concepts such as superordinates, which is enough to create a priming effect in semantic categorization tasks. In our experiments however, a deeper semantic processing was necessary. For example, in the semantic field of birds, the SC voice pronounced: *corbeau, rossignol, cage, voler, nid* ("craw, nightingale, cage, fly, nest") and *PIGEON* ("pigeon") was the target word. In a condition of high intelligibility, these words primed the target word; however, with decreased intelligibility, if the participants only heard the word "cage," this word may have activated its superordinate (e.g., "object") but not its associates such as "bird." Consequently, it may not have primed "pigeon." The absence of a semantic priming effect in addition to the decrease in intelligibility seems to show that greater difficulty in processing auditory signals at a superficial level, causes decreased processing at a (higher) linguistic level.

Consistent with our results, and using an auditory masked priming paradigm, Kouider and Dupoux (2005) demonstrated lexical access without semantic priming. In their experiments, prime was a time-compressed word embedded in a masker composed of time reversed and compressed words. They manipulated the compression rate of the prime to vary its intelligibility. If the prime was not intelligible, they found a significant repetition priming effect on words but not non-words suggesting that this effect "involved a lexical activation of abstract word form." In the same condition of prime compression, they did not find any semantic priming effect. This finding would therefore suggest a dissociation between lexical and semantic processing.

COGNITIVE LOAD

Overall our results are compatible with the idea of cognitive load, in which simultaneously processed information and interactions can either under-load or overload the finite amount of processing capacities. As shown by our results, it is more difficult to process items with more voices in the background. This finding

might be partly due to the high perceptual load and because the target item was embedded in backgrounds that could not be completely ignored (Lavie, 1995, 2005). Processing words in the background might have been particularly cognitively effortful and, semantic processing of a word can be delayed and even prevented if participants have to perform an additional task (i.e., high cognitive load, Hohlfeld et al., 2004; Hohlfeld and Sommer, 2005; Van Petten, 2014). In Experiment 3 we argue that some background words were heard but not deeply semantically processed because it was both very demanding and irrelevant to perform the task. In Experiment 2, 2SC/1SI background words were also very difficult to hear and process (as in Experiment 3 1SC/2SI); however, they were semantically processed as revealed by the semantic priming effect observed. As SC voices were more salient in Experiment 2 2SC/1SI than in Experiment 3 1SC/2SI, participants might have heard more related words and one could argue that semantic priming in our study in fact relied on the chance for participants to hear a SC word. Although this hypothesis is interesting, it does not seem sufficient to explain our results. Indeed, according to this hypothesis, a priming effect should have appeared also in Experiment 3 where participants, in line with previous results from the literature (Brungart, 2001; Hoen et al., 2007), recognized SC words presented in the post-test. Given the overall results of our experiments, we argue that if the ratio of SC/total voices was $<1/2$, participants heard some SC words, but, because of high cognitive load (i.e., intelligibility was low and they focused on the target item) these words were not processed sufficiently deeply to lead to semantic priming (a similar dissociation between lexical and semantic processing was also found by Kouider and Dupoux, 2005). However, if the ratio was $>1/2$, SC words were salient, and participants thus, processed them sufficiently to improve their performance. As this interpretation relies on the post-test effect which is quite small, although significant, more experiments should be performed. For example, using only SC voices and degrading intelligibility by adding noise or filtering the signal instead of adding SI voice would be a good way to test this hypothesis in future studies.

Our results suggest that the increased cognitive load necessary to reconstruct the degraded signal reduced available resources for higher-level processes. This claim is consistent with the Effortfulness Hypothesis (Rabbitt, 1968) that states degraded signals require allocating many cognitive resources to formal processes (i.e., orthographic or phonological), leaving less available cognitive resources to perform higher-level processes (e.g., lexical). It has been shown that hearing-impaired participants are less accurate at recalling previously heard final sentence words than their control peers (Pichora-Fuller et al., 1995). The underlying assumption is that for hearing-impaired participants, the auditory signal is highly degraded and therefore demands more cognitive resources to be formally (i.e., phonologically) processed. As studies usually use recall tasks of lists of unrelated words or digits (Surprenant, 1999; Murphy et al., 2000; Wingfield et al., 2005), verbal working memory only relies on the phonological loop (Baddeley and Hitch, 1974; Baddeley, 2000) and only involves formal processes. Our results showed that higher levels of processing such as semantic activation may be specifically impacted by signal degradation.

Recently, semantic and syntactic integration difficulties if a signal is degraded have been reported in the visual modality (Gao et al., 2011, 2012). Experiments conducted by Gao et al. (2011) using visual noise (i.e., pixel's brightness variation) showed that if participants allocate more resources to formal processes, semantic integration is affected. Indeed, after reading an entire text, participants were worse at recalling the main proposition in the noisy condition. Altogether, these findings suggest that the availability of cognitive resources is involved at various levels during language processing. Whereas previous studies have shown that noise impairs memory systems, our study provides evidence that semantic activation is linked to cognitive resources, independently of memory.

CONCLUSION

This study explored the semantic nature of informational masking in a cocktail party situation. The results of three behavioral studies reveal that the emergence of semantic priming effects relies on prime intelligibility and saliency. These findings question the assumption that signal degradation has no effect on speech processing if target signals can be recognized. The results reveal that high-level processes, such as semantic processing, might not be as automatic as previously thought but are subjected to the limits of cognitive resources. Our study also demonstrates how the cocktail party situation can be used to study the automaticity of linguistic processes.

ACKNOWLEDGMENTS

The first author is funded by a PhD grant from Rhône-Alpes region, France. This research was supported by a European Research Council grant to the SpiN project (no. 209234).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fnhum.2014.00878/abstract

REFERENCES

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/S1364-6613(00)01538-2

Baddeley, A., and Hitch, G. (1974). "Working Memory," in *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 8, ed G. H. Bower (New York, NY: Academic Press), 47–89.

Boulenger, V., Hoen, M., Ferragne, E., Pellegrino, F., and Meunier, F. (2010). Real-time lexical competitions during speech-in-speech comprehension. *Speech Commun.* 52, 246–253. doi: 10.1016/j.specom.2009.11.002

Bronkhorst, A. W. (2000). The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acustica* 86, 117–128.

Brouwer, S., Van Engen, K., Calandruccio, L., and Bradlow, A. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. *J. Acoust. Soc. Am.* 131, 1449–1463. doi: 10.1121/1.3675943

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.* 109, 1101–1109. doi: 10.1121/1.1345696

Brungart, D. S., Chang, P. S., Simpson, B. D., and Wang, D. (2006). Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation. *J. Acoust. Soc. Am.* 120, 4007–4018. doi: 10.1121/1.2363929

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple

simultaneous talkers. *J. Acoust. Soc. Am.* 110, 2527–2538. doi: 10.1121/1.1408946

Cherry, C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229

Collins, A., and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychol. Rev.* 82, 407–428. doi: 10.1037/0033-295X.82.6.407

Collins, A., and Quillian, M. R. (1969). Retrieval time from semantic memory. *J. Verbal Learn. Verbal Behav.* 9, 240–247. doi: 10.1016/S0022-5371(69)80069-1

Cooke, M., Garcia Lecumberri, M. L., and Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *J. Acoust. Soc. Am.* 123, 414–427. doi: 10.1121/1.2804952

Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., et al. (1998). Imaging unconscious semantic priming. *Nature* 395, 597–600. doi: 10.1038/26967

Donnenwerth-Nolan, S., Tanenhaus, M. K., and Seidenberg, M. S. (1981). Multiple code activation in word recognition: evidence from rhyme monitoring. *J. Exp. Psychol. Hum. Learn.* 7, 170–180. doi: 10.1037/0278-7393.7.3.170

Drullman, R., and Bronkhorst, A. W. (2000). Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation. *J. Acoust. Soc. Am.* 107, 2224–2235. doi: 10.1121/1.428503

Dupoux, E., Kouider, S., and Mehler, J. (2003). Lexical access without attention? Explorations using dichotic priming. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 172–184. doi: 10.1037/0096-1523.29.1.172

Durlach, N. (2006). Auditory masking: need for improved conceptual structure. *J. Acoust. Soc. Am.* 120, 1787–1790. doi: 10.1121/1.2335426

Eich, E. (1984). Memory for unattended events: remembering with and without awareness. *Mem. Cognit.* 12, 105–111. doi: 10.3758/BF03198423

Festen, J. M., and Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 88, 1725–1736. doi: 10.1121/1.400247

Forster, K. I., and Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 680–698. doi: 10.1037/0278-7393.10.4.680

Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J. Acoust. Soc. Am.* 115, 2246–2256. doi: 10.1121/1.1689343

Gao, X., Levinthal, B. R., and Stine-Morrow, E. A. (2012). The effects of ageing and visual noise on conceptual integration during sentence reading. *Q. J. Exp. Psychol. (Hove)* 65, 1833–1847. doi: 10.1080/17470218.2012.674146

Gao, X., Stine-Morrow, E. A., Noh, S. R., and Eskew, R. T. Jr. (2011). Visual noise disrupts conceptual integration in reading. *Psychon. Bull. Rev.* 18, 83–88. doi: 10.3758/s13423-010-0014-4

Gautreau, A., Hoen, M., and Meunier, F. (2013). Let's all speak together! Exploring the masking effects of various languages on spoken word identification in multilinguistic babble. *PLoS ONE* 8:e65668. doi: 10.1371/journal.pone.0065668

Grainger, J., and Holcomb, P. J. (2009). Watching the word go by: on the time-course of component processes in visual word recognition. *Lang. Linguist. Compass* 3, 128–156. doi: 10.1111/j.1749-818X.2008.00121.x

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J. Acoust. Soc. Am.* 115, 833–843. doi: 10.1121/1.1639908

Hoen, M., Meunier, F., Grataloup, C.-L., Pellegrino, F., Grimault, N., Perrin, F., et al. (2007). Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. *Speech Commun.* 49, 905–916. doi: 10.1016/j.specom.2007.05.008

Hohlfeld, A., Sangals, J., and Sommer, W. (2004). Effects of additional tasks on language perception: an event-related brain potential investigation. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 1012–1025. doi: 10.1037/0278-7393.30.5.1012

Hohlfeld, A., and Sommer, W. (2005). Semantic processing of unattended meaning is modulated by additional task load: evidence from electrophysiology. *Brain Res. Cogn. Brain Res.* 24, 500–512. doi: 10.1016/j.cogbrainres.2005.03.001

Klauer, K. C., Eder, A. B., Greenwald, A. G., and Abrams, R. L. (2007). Priming of semantic classifications by novel subliminal prime words. *Conscious. Cogn.* 16, 63–83. doi: 10.1016/j.concog.2005.12.002

Kouider, S., and Dehaene, S. (2009). Subliminal number priming within and across the visual and auditory modalities. *Exp. Psychol.* 56, 418–433. doi: 10.1027/1618-3169.56.6.418

Kouider, S., and Dupoux, E. (2005). Subliminal speech priming. *Psychol. Sci.* 16, 617–625. doi: 10.1111/j.1467-9280.2005.01584.x

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 451–468. doi: 10.1037/0096-1523.21.3.451

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends Cogn. Sci.* 9, 75–82. doi: 10.1016/j.tics.2004.12.004

Lewis, J. L. (1970). Semantic processing of unattended messages using dichotic listening. *J. Exp. Psychol.* 85, 225–228. doi: 10.1037/h0029518

Marslen-Wilson, W., and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cogn. Psychol.* 10, 29–63. doi: 10.1016/0010-0285(78)90018-X

Mattys, S. L., Brooks, J., and Cooke, M. (2009). Recognizing speech under a processing load: dissociating energetic from informational factors. *Cogn. Psychol.* 59, 203–243. doi: 10.1016/j.cogpsych.2009.04.001

Mattys, S. L., and Wiget, L. (2011). Effects of cognitive load on speech recognition. *J. Mem. Lang.* 65, 145–160. doi: 10.1016/j.jml.2011.04.004

McClelland, J. L., and Elman, J. L. (1986). The TRACE model of speech perception. *Cogn. Psychol.* 18, 1–86. doi: 10.1016/0010-0285(86)90015-0

McDermott, J. H. (2009). The cocktail party problem. *Curr. Biol.* 19, R1024–R1027. doi: 10.1016/j.cub.2009.09.005

Meyer, D. E., and Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *J. Exp. Psychol.* 90, 227–234. doi: 10.1037/h0031564

Murphy, D. R., Craik, F. I., Li, K. Z., and Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychol. Aging* 15, 323–334. doi: 10.1037/0882-7974.15.2.323

Naccache, L., and Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition* 80, 215–229. doi: 10.1016/S0010-0277(00)00139-6

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of Inhibitionless spreading activation limited-capacity attention. *J. Exp. Psychol. Gen.* 106, 226–254. doi: 10.1037/0096-3445.106.3.226

New, B., Pallier, C., Ferrand, L., and Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique* 101, 447–462. Available online at: http://www.lexique.org

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282

Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *Q. J. Exp. Psychol.* 20, 241–248. doi: 10.1080/14640746808400158

Radeau, M. (1983). Semantic priming between spoken words in adults and children. *Can. J. Psychol.* 37, 547–556. doi: 10.1037/h0080756

Radeau, M., Besson, M., Fonteneau, E., and Castro, S. L. (1998). Semantic, repetition and rime priming between spoken words: behavioral and electrophysiological evidence. *Biol. Psychol.* 48, 183–204. doi: 10.1016/S0301-0511(98)00012-X

Schacter, D. L., and Church, B. A. (1992). Auditory priming: implicit and explicit memory for words and voices. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 915–930. doi: 10.1037/0278-7393.18.5.915

Simpson, S. A., and Cooke, M. (2005). Consonant identification in N-talker babble is a nonmonotonic function of N. *J. Acoust. Soc. Am.* 118, 2775–2778. doi: 10.1121/1.2062650

Spruyt, A., De Houwer, J., Everaert, T., and Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition* 122, 91–95. doi: 10.1016/j.cognition.2011.08.017

Surprenant, A. (1999). The effect of noise on memory for spoken syllables. *Int. J. Psychol.* 34, 328–333. doi: 10.1080/002075999399648

Van den Bussche, E., Van den Noortgate, W., and Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychol. Bull.* 135, 452–477. doi: 10.1037/a0015329

Van Engen, K., and Bradlow, A. (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *J. Acoust. Soc. Am.* 121, 519–526. doi: 10.1121/1.2400666

Van Petten, C. (2014). "Selective attention, processing load, and semantics: insights from human eletrophysiology," in *Cognitive Electrophysiology of Attention*, ed G. R. Mangun (New York, NY: Hardbound), 236–253.

Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in Older Adulthood. What it is and howit interacts with cognitive performances. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Wood, N. L., Stadler, M. A., and Cowan, N. (1997). Is there implicit memory without attention? A reexamination of task demands in Eich's (1984) procedure. *Mem. Cognit.* 25, 772–779. doi: 10.3758/BF03211320

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 20 March 2014; accepted: 12 October 2014; published online: 31 October 2014.

Citation: Dekerle M, Boulenger V, Hoen M and Meunier F (2014) Multi-talker background and semantic priming effect. Front. Hum. Neurosci. ***8****:878. doi: 10.3389/fnhum.2014.00878*

This article was submitted to the journal Frontiers in Human Neuroscience.