



**HAL**  
open science

## Neural correlates of acoustic and semantic cues during speech segmentation in French

Maria del Mar Cordero, Ambre Denis-Noël, Elsa Spinelli, Fanny Meunier

► **To cite this version:**

Maria del Mar Cordero, Ambre Denis-Noël, Elsa Spinelli, Fanny Meunier. Neural correlates of acoustic and semantic cues during speech segmentation in French. Interspeech 2022, Sep 2022, Incheon, South Korea. pp.4058-4062, 10.21437/Interspeech.2022-10986 . hal-03916475

**HAL Id: hal-03916475**

**<https://hal.science/hal-03916475>**

Submitted on 30 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Neural correlates of acoustic and semantic cues during speech segmentation in French

Maria del Mar Cordero<sup>1</sup>, Ambre Denis-Noël<sup>2</sup>, Elsa Spinelli<sup>3</sup>, Fanny Meunier<sup>1</sup>

<sup>1</sup>Université Côte d'Azur, CNRS, BCL, France

<sup>2</sup>Complexity and Cognition Laboratory, Université Côte d'Azur, Nice, France

<sup>3</sup>Université Grenoble Alpes, CNRS, LPNC, Grenoble, France

maria-del-mar.cordero-rull@etu.univ-cotedazur.fr, ambre.denis-noel@univ-cotedazur.fr, elsa.spinelli@univ-grenoble-alpes.fr, fanny.meunier@unice.fr

## Abstract

Natural speech is highly complex and variable. Particularly, spoken language, in contrast to written language, has no clear word boundaries. Adult listeners can exploit different types of information to segment the continuous stream such as acoustic and semantic information. However, the weight of these cues, when co-occurring, remains to be determined. Behavioural tasks are not conclusive on this point as they focus participants' attention on certain sources of information, thus biasing the results. Here, we looked at the processing of homophonic utterances such as *l'amie* vs *la mie* (both /lami/) which include fine acoustic differences and for which the meaning changes depending on segmentation. To examine the perceptual resolution of such ambiguities when semantic information is available, we measured the online processing of sentences containing such sequences in an ERP experiment involving no active task. In a congruent context, semantic information matched the acoustic signal of the word *amie*, while, in the incongruent condition, the semantic information carried by the sentence and the acoustic signal were leading to different lexical candidates. No clear neural markers for the use of acoustic cues were found. Our results suggest a preponderant weight of semantic information over acoustic information during natural spoken sentence processing.

**Index Terms:** spoken speech recognition, speech segmentation, intonational cues, coarticulation, French, ambiguity

## 1. Introduction

Speech is a continuous signal that is rarely marked by pauses between words, therefore it is unclear how listeners effortlessly solve the task of converting continuous streams of phonemes into separate and linguistically meaningful units (a process known as segmentation). As shown in previous studies, a wide range of cues can be used by listeners to locate word boundaries, varying from suprasegmental features (1,2) to phonotactics (3,4). However, listeners seem to resort to different strategies for encoding word boundaries depending on their native language. For example, English and Dutch speakers rely acoustic sources of information such as durational differences and stress patterns (5–8). However, these cues are non-contrastive for French speakers, and it is much less clear which strategy they rely upon (9,10).

In French, the syllable plays an important role in speech segmentation (4,11–13). However, sandhi phenomena in French such as *liaison* and *elision*, lead to resyllabification,

making the syllabic strategy inadequate as it could lead to the identification of an inappropriate lexical candidate. As an example, the spoken sequence /lami/ matches two different lexical representations: *l'amie* ('the friend') and *la mie* ('the crumb'), yet behavioral experiments have shown that, in isolation, these sequences are identified and discriminated correctly at 75% (14). This seems to be achieved thanks to fine acoustic cues that, without being perfectly robust, allow for scores well above chance. In a cross-modal priming study, Spinelli et al. (15) observed facilitatory effects of acoustic features in pivotal /n/ when targets matched the speaker's intended segmentation for homophonic sequences such as *un air* (/ɛ̃nɛʁ/ "a tune", target: AIR) and *un nerf* (/ɛ̃nɛʁ/ "a nerve", target: NERF). However, such effects were still observed, though weaker, when targets mismatched the intended segmentation. These results showed that resyllabification does not hinder word recognition. Acoustic analyses of this type of sequences showed fine differences in duration (resyllabified consonants were overall of shorter duration) and revealed an F0 rise at the beginning of the content word (see (14) for other differences in formant structure). Taken all together, while it appears that some fine acoustic cues can be exploited to a certain extent to segment speech, it is not clear what their weight is when other information, such as semantics, is available.

Research on the use of multiple cues seems to support the view that the role of acoustic cues in natural speech processing may be overestimated. In particular, previous work has demonstrated that the use of cues may depend on their availability (16,17) but also on the attentional focus required during the task (e.g.,(18)). In a series of experiments, Mattys et al. (16) presented syllables extracted from real words placed after either intact or noisy contexts. Contexts were conformed of either a word or a nonword. Results showed that acoustic cues were only used when context information was not optimal (noisy), and hindered the extraction of information at lexical or sentence level in order to achieve correct segmentation. Later, Mattys and Melhorn (18) tested near-homophones like *plum pie/plump eye* in embedded in a sentence. Here, listeners relied on semantic information when the context was incongruent. However, when the task required attention to be directed to the acoustics of the utterances, effects reflected a shift towards reliance upon acoustic information over lexical to segment speech. While these studies have clearly demonstrated the possibility of using different types of cues to segment speech, they further suggested a crucial role of the task participants are asked to perform. Effects observed in such studies show that the interplay between lower- and higher-level cues in optimal situations is not yet clear. To further explore the use of different

cues for speech segmentation, in this study we tested the use of acoustic cues present in homophonic utterances such as *la mie/l'amie* embedded in sentences, i.e. complex contexts.

The importance of the methodological approach in researching speech processing and segmentation is well established. To date, many experiments have focused on behavioural studies presenting a task to participants while processing either speech in noise (e.g.,(16)) or natural speech (e.g.,(19)). It appears that manipulating acoustic information influences the listener's comprehension (15,20). Nevertheless, behavioural measures only provide insight on the post-segmentation processing rather than the ongoing segmentation process. Furthermore, it is not clear from behavioral data whether multiple cues are modulated during ecological speech perception. To assess whether acoustic cues are relevant for segmenting speech in French by overcoming the task issue, Do Carmo Blanco et al. (21) carried a mismatch negativity (MMN) experiment using a modified oddball paradigm. For this paradigm, a series of frequent sounds (standards) are presented intermingled with a rare stimulus (deviant) that is presented in a lower rate. MMN allows to explore the electrophysiological correlates to both relevant and irrelevant fine prosodic and phonemic cues even without focused attention (e.g.,(22)). Exploitation of non-contrastive cues present in the sequences was tested for both syllable (/la/) and utterance (/lami/) productions, i.e., both at the acoustic level and in lexical cues. Results showed clearly that acoustic features present in /lami/ are regular and/or salient enough for the listener's brain to extract that information and associate it with the intended segmentations without attentional focus.

In the present experiment, we aimed to build on these results and test the same homophonic sequences with addition of a semantic context to see which cues would be considered. We thus looked at the semantic integration of ambiguous sequences such as *l'ami/la mie*. Evidence from event-related brain potentials (ERPs) have shown that semantically anomalous words to a given context yield a negative wave ERP peaking at about 400 ms post-stimulus, known as N400 (23) (N400 being associated to the process of integrating semantic information). Its amplitude is related to the difficulty with which the speaker processes semantic information when it is anomalous and, therefore, the possible difficulty in accessing the lexical representation (e.g.,(24)). Some studies have addressed the interplay between processing early acoustic and late semantic information (25,26). For example, Aydelott et al. (26) examined the effects on semantic processing when presented with degraded acoustic features, where sentence-final words were presented in either congruent or incongruent contexts (contexts were either low-pass filtered or intact). Results showed that acoustic degradation modulates the N400 amplitude, which is lower for incongruent targets in filtered contexts. This showed that the acoustic manipulation reduced the availability of semantic information, especially in the incongruent context. In this study, we want to explore the impact on semantic integration of intact acoustic features leading to different segmentations.

In the present study we tested whether acoustic characteristics present in homophonic sequences like /lami/ are robust or salient enough for the native listener when presented in a sentence semantically oriented, and, if so, whether and how they influence lexical access and, as a result, speech comprehension. As ERP measures provide a direct response during processing of sentences containing homophonic sequences, the modulation of the N400 component was explored to investigate the influence of acoustic and semantic

features on the listeners' responses. This analysis was applied using three conditions. First, non-homophonic, semantically anomalous sentences were generated to yield larger N400 responses as control condition. Then, two experimental conditions were set up: one where the semantic information carried by the sentence matched the acoustic information of the ambiguous sequence; and another in which the two sources of information were incongruent. We hypothesized that congruent and control conditions will show distinctly different N400 effects due to the nature of the sentences. In addition, the incongruent condition will elucidate larger N400 effects than the congruent one if acoustic information is factored in as a cue when semantic information is also available. Otherwise, if acoustic information is not considered by the listener, incongruent and congruent conditions will reveal no differences in their N400 amplitudes, and incongruent sentences will elucidate smaller N400 effects compared to control sentences.

## 2. Method

### 2.1. Participants

Twenty-three participants (13 females, mean age 20, in range 18-26) were rewarded at the end of the experiment for their participation. All were native French speakers, right-handed, and all had normal to corrected vision. None reported any psychiatric or neurological disorders or current use of medications thought to affect the central nervous system.

### 2.2. Materials

We used 24 pairs of phonologically identical nominal sequences in French (e.g., *l'amie/la mie*, both pronounced /lami/). Each word was contained in sentences with a strong semantic constraint in favour of one of the candidates, for example *Le collégien a rencontré l'amie de Paul* ("The student met the friend of Paul") or *Les édentés apprécient la mie de pain* ("The toothless like the crumb of bread"). Sequences had the following structure: i) Determiner *la* + the consonant-initial candidate –e.g., *la mie* ("the crumb")–; ii) determiner *l'* + the vowel-initial candidate –e.g., *l'amie* ("the friend")–.

Sentences were recorded by a French speaker, naïve to the study, normalized at 65 dB-A and to 44.1 kHz. There were four sentences per pair to create a 2x2 design. To generate the different context conditions without introducing acoustic variations, two acoustic manipulations were performed using the Audition software (see Figure 1). *Cross-splicing* (panels A on the left) consisted in exchanging the homophone target sequences, which yielded sentences with contexts that were not congruent with the homophone candidate included (Congruent context condition). *Identity-splicing* (panels B on the right) consisted in exchanging a target sequence with another production of the same nominal sequence; congruent sentences were generated after the identity-splicing (Incongruent context condition). This amounts to a total of 48 sentences for each condition.

Forty-eight additional sentences were included as Control condition. To control for the effect of contextual incongruity and allow comparisons in the analysis of EEG data, Control sentences contained a semantically non-congruent, non-homophonic nominal sequence with the same structure as the other conditions: i) half *la* + consonant-initial noun (such as *Le cordonnier répare la tendresse du directeur* "The shoemaker repairs the tenderness of the director"); ii) the other half *l'* + vowel-initial noun (*Le chien tire sur l'amidon pendant la promenade* "The dog pulls on the starch during the walk"). The same acoustic

manipulation was performed to create Control sentences, a cross-splicing on the non-homophonic nominal sequences.

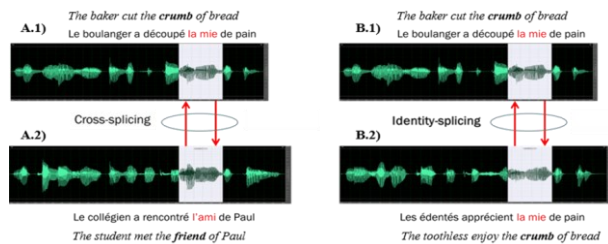


Figure 1: Illustration of the acoustic manipulations performed to generate the stimuli: A) cross-splicing, B) identity-splicing.

Written word frequency and spoken word frequency for each homophone candidate are shown in Table 1 below. Such measures were extracted from the published corpus *Lexique* (27,28). The Cloze Probability of the sentences was calculated from the responses of 32 participants who were asked to complete a gap-filling task starting from the word of interest. Participants had to complete sentences from the two homophone candidates in only one of the two contexts. Cloze probability ratios were significantly different by context condition ( $F(3, 92) = 6.22, p < 0.0007$ ). Nevertheless, we did not find significant differences by homophone target across context conditions ( $F(3, 23) = 0.85, p = 0.66$ ).

Table 1: Mean (SD) of frequencies, length, and position of homophones for both item types (vowel and consonant initial), and both context types (congruent and incongruent).

	Written freq.	Spoken freq.	Word length (s)	Sentence pos. (s)
Vowel initial	22.90 (30.43)	27.54 (46.30)	0.640 (0.126)	1.816 (0.456)
Consonant init.	52.01 (167.60)	57.52 (199.02)	0.631 (0.137)	1.587 (0.352)
Congruent	37.57 (122.65)	42.90 (146.87)	0.639 (0.139)	1.739 (0.441)
Incongruent	37.54 (121.37)	42.61 (145.35)	0.635 (0.124)	1.657 (0.396)

We assessed for differences among durations of critical words through a one-way Anova, showing no significant differences for items across conditions ( $F(2,141) = 10.36, p = 3.06$ ). Sentence length was kept between 6-13 words ( $M = 8.54 \pm 1.48$ ). A one-way Anova revealed no significant differences for sentence duration across conditions ( $F(2,141) = 1.50, p = 3.06$ ). Position of the target word can lead to variations in tested effects (29–31); a one-way Anova revealed no significant differences ( $F(2,141) = 1.82, p = 3.06$ ) for relative positions of words of interest within sentences across conditions.

### 2.3. Procedure

The experiment took place in an electrically and acoustically shielded booth, sitting in a comfortable armchair in front of an LCD. After participants had read and signed all the forms, they were instructed to listen attentively to French sentences. While a sentence was played, a fixation cross was displayed in the

centre of the screen to keep the participants' sight focused and minimize muscle artifacts (32). In a pseudorandomized order, a written sentence appeared on the screen after some spoken sentences were played. After each written sentence, participants had to indicate on a two-alternative forced choice whether the written sentence was semantically related to the last sentence they have heard. This sentence judgment task was used as an attentional control. The number of sentences prior to the task was randomized to prevent from anticipatory effects. This setting was repeated along 5 blocks: 1 training block consisting of 11 spoken sentences (4 incongruent, 4 congruent, 3 control), after what followed 4 testing blocks each containing 12 congruent sentences, 12 incongruent, and 15 control. All stimuli sets were randomized among participants, keeping the training stimuli consistent but displayed in a different order. To avoid proximity of homophone candidates, they were split into four different experimental blocks, such that every block contained a candidate of each condition (see Table 2 as an example of the pair *l'amie-la mie*). Following the sentences in Table 2, in blocks 1 and 2 *mie* was the target homophone, whereas it was *amie* for blocks 3 and 4.

Table 2: Pair *l'amie - la mie* as example of the distribution of sentences along the experiment.

Block	Sentence (condition)
Block 1	La famille apprécie <u>la mie</u> du marié. <i>The family appreciates <u>the crumb</u> of the groom.</i> ( <i>La mie</i> incongruent)
	Les édentés apprécient <u>la mie</u> de pain. <i>The toothless appreciate <u>the crumb</u> of bread.</i> ( <i>La mie</i> congruent)
Block 3	Le boulanger a découpé <u>l'amie</u> de pain. <i>The baker cut out <u>the friend</u> of bread.</i> ( <i>L'amie</i> incongruent)
	Le collégien a rencontré <u>l'amie</u> de Paul. <i>The student met <u>the friend</u> of Paul.</i> ( <i>L'amie</i> congruent)

### 2.4. EEG recording and analysis

Neurophysiological measures were acquired using the Curry 8.0 XS software (Compumedics, Neuroscan SynAmps 2/RT, 24-bit sampling method), version 8.0.6 X. EEG was recorded from 64 active channels, mounted in an elastic cap (Neuroscan Quik-Cap Neo Net, with 4 integrated bipolar leads for vertical and horizontal eye movements, fitted according to the standard international 10/20 positioning system). Recording bandwidth was 0.1–400 Hz and sampling rate 1000 Hz. EEG sessions were recorded on average for 22.02 s ( $\pm 1.06$ ). Impedances were kept below 5 k $\Omega$ . No online filter was used.

Eight pairs of sequences were discarded due to technical issues and 2 participants were rejected due to a high rejection of trials. EEG data were preprocessed and analysed in EEGLAB (v2019). Auditory evoked potentials to the critical sequence (/lami/) were analysed over a 1300 ms window, including a prestimulus baseline of 200 ms, after being offline filtered between 0.1 and 48 Hz. Mean number of trials per participant was  $99.68 \pm 6.68$ . Data was decomposed using a preconditioned independent component analysis for real data (Picard, (33), followed by an automatic independent component

classifier (ICLabel, (34)) to correct for eye, muscle, and channel noise artifacts (>80%). Interpolation and re-referencing to average were performed as a final step to optimize signal-to-noise ratio (STNR). We run an (spherical) interpolation for missing channels and the reference was recomputed to average.

To obtain the large means of each experimental condition, epochs were averaged over subjects and then split by conditions. Due to the noisy nature of EEG signal during auditory processing and to the noise present in the baseline period of epochs, we performed an analysis based on global field potential (GFP, see Figure 2). GFP corresponds to the spatial root mean squared of amplitude values at all electrodes (35) and is calculated from the grand average across all subjects and all experimental conditions. Selected time windows were chosen according to periods exhibiting a change over the whole recording, indicating a variation of the evoked potential fields over time. Groups of electrodes displaying significant effects of interest were determined via permutation tests (36). Signal amplitude was analysed through permutation tests carried in each time window separately using the R package *lmPerm*.

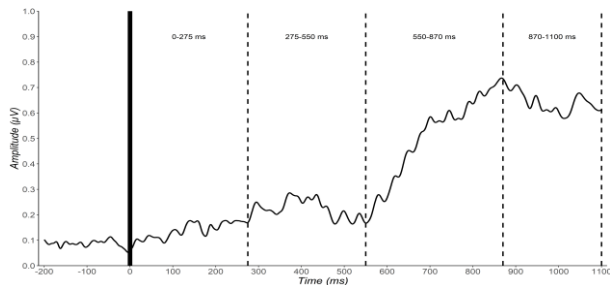


Figure 2: GFP waveform calculated 1100 ms following critical stimulus onset (/lami/). Dashed lines showing time windows presenting a change of voltage.

### 3. ERP results

ERP waveforms time-locked to the stimulus onset are presented in Figure 3. Due to space limits, scalp maps are not included.

*Early positivity (275-550 ms).* Anterior scalp locations presented a positive polarity, more prominent for congruent compared to other context types. Permutation tests revealed a significant effect for congruent compared to control contexts over the time window of 345-550 ms.

*N400 (550-870 ms, 870-1100 ms).* Topographic maps showed a negativity spread over frontocentral electrodes slightly right lateralized for control condition. This negative amplitude reached statistical significance ( $p < 0.025$ ) at 745-870 ms over F5, FC5, and FC3 electrodes only for control compared to congruent conditions. Starting at 870 ms, the negative-going amplitude reached significance for frontocentral electrodes (Cz, C2, C4, FC2, FC4) for control compared to incongruent conditions ( $p < 0.025$ ).

### 4. Discussion and Conclusion

The present study was designed to measure the impact of acoustic and semantic information on speech segmentation. Homophonic sequences were embedded in congruent contexts, favouring the meaning of the auditory candidate (*l'amie*), or incongruent contexts, favouring the meaning of the other candidate (*la mie*). Our data show direct processing of spoken speech and demonstrate the relevance of context over acoustic

information during segmentation of ambiguous utterances in ecological speech settings.

No amplitude differences among context types were observed in N1-P2 components, suggesting that early ERP components sensitive to physical attributes of auditory input remain unaffected by the F0 features present in the homophonic sequences. This is in line with previous studies showing that when processing continuous speech, early components elicited much smaller amplitudes (2) or even no onset effects (37,38). Including more trials may increase the possibility to visualize small amplitude components like N1-P2. Control conditions yielded larger N400 amplitudes after onset of non-homophonic utterances as a response to incongruous words. In contrast, no effect was found for homophonic candidates within unfavourable contexts compared to candidates within favourable ones. This suggests that without focused attention acoustic cues are not used to segment speech when contextual information is available. This would indicate that although listeners exploit the acoustic attributes of homophonic utterances to access the intended segmentation (e.g.,(39)), when semantic information is available it serves as a more reliable or salient source and is extracted by the listeners' perceptual system. This is related to previous findings showing that the semantic N400 is more robust than the phonological auditory N400. Perrin and Garcia-Larrea (40) showed that when no task was performed, unrelated and related words did not elicit N400 effects.

We demonstrated that a similar response is elicited when listeners are exposed to non-contrastive cues present in /lami/. Our results showed that there are no clear cues to the use of acoustics when speech is processed unconsciously, which did not lead to confusion nor impede understanding. Presenting listeners with spoken sentences without requiring their attention to be involved, allowed us to assess the neural correlates of processing co-occurring cues in an unattended and unconscious manner. No differences were found between our congruent and incongruent contexts, leading us to believe that diverging cues have no impact on sentence comprehension. Contrary to what was found in the MMN study, our results suggest that when semantic information is present, acoustic cues are less salient to listeners and do not rely on such information to achieve the ultimate goal, comprehension. This demonstrates that higher-order information does not appear to alter the low-level analysis of conflicting acoustic signals to avoid segmentation errors.

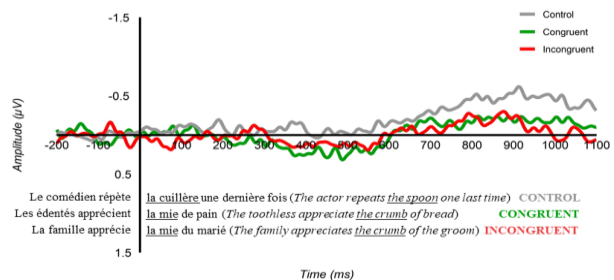


Figure 3: ERP waveform averaged over electrodes presenting statistical significance effects.

### 5. Acknowledgements

This work was supported by the French government (UCA-Jedi), project managed by the National Research Agency (ANR-15-IDEX-01) and by the CNRS Service and Research Unit MSHS Sud-Est, USR 3566, of the Université Côte d'Azur.

## 6. References

- [1] Quené H. Durational cues for word segmentation In Dutch. *J Phon.* 1992 Jul 1 (20), 331-350.
- [2] Sanders LD, Neville HJ. An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cogn Brain Res.* 2003 Feb 1;15(3):228-40.
- [3] McQueen JM. Segmentation of Continuous Speech Using Phonotactics. *J Mem Lang.* 1998 Jul 1;39(1):21-46.
- [4] Mehler J, Dommergues J, Frauenfelder U, Segui J. The syllable's role in speech segmentation. *J Verbal Learn Verbal Behav.* 1981 Jun 1;20:298-305.
- [5] Altenberg EP. The perception of word boundaries in a second language. *Second Lang Res.* 2005 Oct 1;21(4):325-58.
- [6] Ito K, Strange W. Perception of allophonic cues to English word boundaries by Japanese second language learners of English. *J Acoust Soc Am.* 2009 Apr;125(4):2348-60.
- [7] Shatzman K, McQueen J. The modulation of lexical competition by segment duration. *Psychon Bull Rev.* 2007 Jan 1;13:966-71.
- [8] Zora H, Schwarz I-C, Heldner M. Neural correlates of lexical stress: mismatch negativity reflects fundamental frequency and intensity. *NeuroReport.* 2015 Sep 9;26(13):791-6.
- [9] Tyler MD, Cutler A. Cross-language differences in cue use for speech segmentation. *J Acoust Soc Am.* 2009 Jul;126(1):367-76.
- [10] Welby P. The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Commun.* 2007 Jan;49(1):28-48.
- [11] Content A, Keams R, Frauenfelder U. Boundaries versus Onsets in Syllabic Segmentation. *J Mem Lang.* 2001 Aug 1;45:177-99.
- [12] Dumay N, Frauenfelder U, Content A. The Role of the Syllable in Lexical Segmentation in French: Word-Spotting Data. *Brain Lang.* 2002 Apr 1;81:144-61.
- [13] Cutler A, Butterfield S. Rhythmic cues to speech segmentation: Evidence from juncture misperception. *J Mem Lang.* 1992 Apr 1;31(2):218-36.
- [14] Spinelli E, Welby P, Schaegis A-L. Fine-grained access to targets and competitors in phonemically identical spoken sequences: the case of French elision. *Lang Cogn Process.* 2007 Sep 1;22(6):828-59.
- [15] Spinelli E, McQueen JM, Cutler A. Processing resyllabified words in French. *J Mem Lang.* 2003 Feb 1;48(2):233-54.
- [16] Mattys SL, White L, Melhorn JF. Integration of multiple speech segmentation cues: a hierarchical framework. *J Exp Psychol Gen.* 2005 Nov;134(4):477-500.
- [17] Sanders LD, Neville HJ. Lexical, Syntactic, and Stress-Pattern Cues for Speech Segmentation. *J Speech Lang Hear Res.* 2000 Dec;43(6):1301-21.
- [18] Mattys SL, Melhorn JF. Sentential, lexical, acoustic effects on perception of word boundaries. *J Acoust Soc Am.* 2007 Jul;122(1):554-67.
- [19] van Donselaar W, Koster M, Cutler A. Exploring the Role of Lexical stress in Lexical Recognition. *Q J Exp Psychol Sect A.* 2005 Feb;58(2):251-73.
- [20] Shoemaker E. Durational cues to word recognition in spoken French. *Appl Psycholinguist.* 2014 Mar;35(2):243-73.
- [21] Do Carmo-Blanco N, Hoen M, Pota S, Spinelli E, Meunier F. Processing of non-contrastive subphonemic features in French homophonous utterances: An MMN study. *J Neurolinguistics.* 2019 Nov;52:100849.
- [22] Näätänen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clin Neurophysiol.* 2007 Dec 1;118(12):2544-90.
- [23] Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science.* 1980 Jan 11;207(4427):203-5.
- [24] Kutas M, Federmeier KD. Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annu Rev Psychol.* 2011;62:621-47.
- [25] Sivonen P, Maess B, Lattner S, Friederici A. Phonemic restoration in a sentence context: Evidence from early and late ERP effects. *Brain Res.* 2006 Dec 1;1121:177-89.
- [26] Aydelott J, Dick F, Mills DL. Effects of acoustic distortion and semantic context on event-related potentials to spoken words. *Psychophysiology.* 2006;43(5):454-64.
- [27] New B, Pallier C, Brysbaert M, Ferrand L. Lexique 2 : A new French lexical database. *Behav Res Methods Instrum Comput.* 2004 Aug;36(3):516-24.
- [28] New B, Pallier C, Ferrand L. Une base de données lexicales du français contemporain sur internet : LEXIQUE™. 2001:21.
- [29] Adda-Decker M, Gendrot C, Nguyen N. Contributions du traitement automatique de la parole à l'étude des voyelles orales du français. 2008 Jan 1;
- [30] Meunier C, Espesser R. Vowel reduction in conversational speech in French: The role of lexical factors. *J Phon.* 2011 Jul;39(3):271-8.
- [31] Ernestus M, Lahey M, Verhees F, Baayen RH, Jasa E. Lexical frequency & voice assimilation. *J Acoust Soc Am.* 2006;120:1040-51.
- [32] Cohen MX. *Analyzing Neural Time Series Data: Theory and Practice.* Cambridge, MA, USA: MIT Press; 2014. 600 p.
- [33] Ablin P, Cardoso J-F, Gramfort A. Faster Independent Component Analysis by Preconditioning With Hessian Approximations. *IEEE Trans Signal Process.* 2018 Jun 4;66.
- [34] Pion-Tonachini L, Kreutz-Delgado K, Makeig S. ICLabel: An automated electroencephalographic independent component classifier, dataset, website. *NeuroImage.* 2019 Sep;198:181-97.
- [35] Lehmann D, Skrandies W. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr Clin Neurophysiol.* 1980 Jun;48(6):609-21.
- [36] Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods.* 2007 Sep 1;164:177-90.
- [37] Friederici A, Pfeifer E, Hahne A. Event-related brain potentials during natural speech processing; Effect of semantic morphological and syntactic violations. *Brain Res Cogn Brain Res.* 1993 Nov 1;1:183-92.
- [38] Hagoort P, Brown CM. ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia.* 2000;38(11):1531-49.
- [39] Spinelli E, Grimault N, Meunier F, Welby P. An intonational cue to word segmentation in phonemically identical sequences. *Atten Percept Psychophys.* 2010 Apr;72(3):775-87.
- [40] Perrin F, García-Larrea L. Modulation of the N400 potential during auditory phonological/semantic interaction. *Cogn Brain Res.* 2003 Jun;17(1):36-47.