



HAL
open science

The N400 reveals implicit accent-induced prejudice

Maud Pélissier, Emmanuel Ferragne

► **To cite this version:**

Maud Pélissier, Emmanuel Ferragne. The N400 reveals implicit accent-induced prejudice. *Speech Communication*, 2022, 137, pp.114-126. 10.1016/j.specom.2021.10.004 . hal-03702086

HAL Id: hal-03702086

<https://u-paris.hal.science/hal-03702086v1>

Submitted on 22 Jul 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

The N400 reveals implicit accent-induced prejudice

Maud Pélissier^{a1}, Emmanuel Ferragne^a

^aLaboratoire de Phonétique et Phonologie, UMR 7018, CNRS & Université Sorbonne Nouvelle, 19 rue des Bernardins, 75005 Paris, France

Corresponding author: Maud Pélissier, mpelissier.pro@gmail.com

¹ Present address: Universitetet i Agder, Institutt for fremmedspråk og oversettelse, Universitetsveien 51, 4630 Kristiansand S, Norway

Abstract

Our perception of someone's accent influences our expectations about what they might say or do. In this experiment, EEG data were recorded while participants listened to cliché sentences matching or not the stereotypes associated with the speaker's accent (upper-class Parisian accent or *banlieue* accent, a negatively connoted accent associated with youth from suburban areas; e.g. "I always listen to rap in my car" said with a banlieue accent (congruent) or an upper-class accent (incongruent)). Mismatches between social accent and stereotypical content triggered an event-related potential (ERP) known as the N400, albeit more anterior than the one observed for semantic violations, as well as a P3. These results are in line with other studies – conducted in particular with gender stereotypes – suggesting that stereotypes are stored in semantic categorical knowledge and that mismatches trigger integration difficulties and checking and updating mechanisms, and extend them to socially marked accents.

Keywords: event-related potentials; social stereotypes; accent; N400; prejudice

1. Introduction

When we listen to someone, we immediately and automatically form expectations about what they are going to say, based on their perceived age, gender, geographic origin or social background. Our intuitions are greatly influenced by the person's accent (Munson, 2007; Clopper and Bradlow, 2009; Lev-Ari and Keysar, 2010). A sentence can have a different meaning depending on who says it and in what context: we use what we know about the speaker to interpret their message. Traditional models of linguistic processing postulate that after the identification of phonemes and word forms from phonological information, the message is first processed automatically and unconsciously at the syntactic level and then the semantic one (Friederici, 2002), while pragmatic information (including indexical aspects such as age, gender, geographic origin or social background) is integrated afterwards (Lattner and Friederici, 2003). However, it would make sense for the listener to take pragmatic aspects of the message into account much earlier, especially because i) numerous words do not have any intrinsic meaning without a specific context, and ii) if language has evolved to support social interactions, social/pragmatic information must be processed at a very early stage (van Berkum et al., 2008). Our objectives in this study were twofold: (1) better understand how stereotypical information is processed in real time and integrated into semantic processing, and (2) evaluate the extent to which a socially marked accent in French (banlieue accent) activates negative stereotypes associated with criminality, a prominent feature of the banlieue stereotype.

Electro-encephalography (EEG), and in particular Event-Related Potentials (ERP) analyses, are well-suited to examining the different steps of language processing in real-time, as their temporal resolution is excellent (1 ms in our case). ERPs are changes in the electrical activity of the brain triggered by specific events, such as the processing of a particular word, and that can be recorded at the surface of the scalp by electrodes. Two components have been linked to the integration of stereotypical information in language processing: the N400, a centro-posterior negative deflection typically maximal around 400 ms, and larger after the onset of a semantically incongruous or less predictable word in a sentence context (Kutas and Hillyard, 1980), and the P600, a large parietal positivity peaking between 500 and 900 ms after a syntactic violation, a *garden-path* effect or while processing syntactically complex sentences (e.g., Kaan et al., 2000). A larger N400 is usually interpreted as a sign of difficulties in lexico-semantic processing, while the P600 reflects late, second-stage syntactic processing, and is associated with the cognitive costs of repairing the sentence and of

re-integrating syntactic, but also semantic and stereotypical information (Osterhout et al., 1997; Lattner and Friederici, 2003).

Several studies have used ERPs to show that indexical information about the speaker is in fact processed very early and can be used to automatically trigger expectations during sentence processing. A large body of literature has focused on gender stereotypes. Osterhout et al. (1997) studied responses to reflexive pronouns in English that referred either to a definitionally gendered noun (e.g., “uncle”, “mother”, “maid”), or to a stereotypically gendered one (e.g., “feminist”, “model”), while the gender of the pronoun was either congruent or not with the (definitional or stereotypical) gender of the antecedent (e.g., “The feminist made herself/himself heard”). A mismatch between the gender of the pronoun and the gender activated by the noun led to a P600, showing that a stereotypical association with a particular gender is enough to affect grammatical expectations. The authors note that this response was “quite distinct from the N400 effect elicited by pragmatic anomalies” (i.e., such as “I take my coffee with cream and *dog*”), and not related to (gender) stereotypes (Osterhout, 1997, 278). Canal et al. (2015) and Su et al. (2016) obtained similar results with the same paradigm in English and Mandarin Chinese, respectively.

Other studies have found that a gender stereotype mismatch was followed by an N400 effect but no P600. Molinaro et al. (2016) used nouns stereotypically biased towards a specific gender in Spanish (e.g., *los mineros*, the miners), and manipulated the congruence of the gender of the noun with stereotypes. When the gender of the noun did not match the associated stereotypes (e.g., *las mineras*, the_[+F] (*female*) miners), a sustained negativity resembling an N400 effect, but no P600, appeared. This can be explained by the fact that, contrary to Osterhout et al. (1997), no grammatical expectations are built, as the gender information is already available on the noun. The mismatch consists in encountering the word ‘*mineras*’ itself, as miners are not expected to be women. Other teams worked with word pairs: in White et al. (2009) or Wang et al. (2017), participants had to evaluate the congruence of a word stereotypically associated with men or women with a gender category prime (“men/male” or “women/female”). They observed a slow-down in response times when the stereotype was incongruent with the prime, as well as a larger N400 effect for incongruous pairs, but no P600 congruency effect (Wang et al., 2017).

All these studies focused on reading: sentences were displayed word by word on a computer screen. Stereotype-related inferences thus stemmed from lexical and world knowledge. However, when we talk with someone, we do not use only the content of their message to interpret their meaning, but we also exploit all available extra-textual information, such as our knowledge of their personality or, for a stranger, their appearance and what they sound like. We are particularly interested in the latter; yet, very few studies have examined how cues that can be extracted from the speaker’s accent (such as, for example, their gender or social and geographic background) create stereotypical expectations and how these in turn can bias the semantic content. We only know of three studies that directly focused on that question, again with gender stereotypes: Grant et al. (2020), Lattner and Friederici (2003), and van Berkum et al. (2008). The last two used sentences containing self-references to the speaker which were either congruous or not with the gender information that could be implicitly extracted from the speaker’s voice, such as a man saying “I bought the same *sewing machine* as Ella”. However, the results of these two studies are contradictory: stereotype mismatches in Lattner and Friederici (2003) were followed by a P600, while those in van Berkum et al. (2008) triggered an N400 effect. Grant et al. (2020) had male and female speakers read statements about stereotypically feminine (fashion) or masculine (sports) topics, which could contain semantic errors congruent or not with gender stereotypes (e.g., women are not expected to be wrong about fashion and utter statements such as “Vogue’s shoe specialist wrote an article about a

suit in the latest edition”). The authors found that the N400 effect of semantic correctness (i.e., the difference between a semantically correct and incorrect statement about sports or fashion) was larger when stimuli were incongruent with gender stereotypes.

While these studies mostly focused on gender stereotypes, there is a large body of literature on the implicit activation of ethnic stereotypes and prejudice using face priming tasks (see Ibáñez et al. (2009) and Amodio (2014) for a review). For example, in the Weapons Identification Task (WIT) (Payne, 2001), participants see a prime consisting in digital photographs of White and Black male faces for 200 ms, followed by targets which are photos of either handguns or hand tools, for another 200 ms. They are asked to classify the target as a gun or a tool by pressing one of two keys. Results show that, when primed by a Black face, White participants identify guns faster and misidentify tools as guns more often than when primed by a White face. In an adjective classification task (Fazio et al., 1995), the prime is also a White or Black face, and the target is a word that has to be classified as either “pleasant” (e.g., attractive or wonderful) or “unpleasant” (e.g., disgusting or annoying). In a similar way to the WIT, there is greater facilitation among White Americans (i.e., words are correctly classified faster) for positive words following White faces and negative words following Black faces. These studies, when using EEG, look at different components than the N400 effect and P600. In particular, several studies have demonstrated a larger amplitude of the P3, an ERP component linked to cognitive reorientation and context updating (Bartholow et al., 2006) or of an error-related negativity (Amodio et al., 2004), when the target word or image does not match stereotypes, and is therefore more unexpected. Amodio et al. (2004) interpret their results as a sign of the resolution of the conflict stemming from the mismatch between the image and what has been implicitly activated by stereotypes. Critically, this effect can be observed even when participants have been informed of the nature of the tasks and therefore make conscious efforts to reduce the effect of implicit prejudice. This activation is almost instantaneous – ethnic or gender information visible on a face can affect ERPs as early as 145 ms after the presentation of that face (see Ibáñez et al. (2009) for a review).

Very few studies have investigated the automatic activation of social or ethnic stereotypes by the accent of the speaker rather than their face - van Berkum et al.’s study (2008) included social class stereotypes, but they were few and integrated to the other categories of pragmatic mismatches (age and gender), and were not analyzed separately. Hansen et al. (2017) showed that an accent can activate ethnic representations extremely fast. In their EEG experiment, participants heard a very short sentence pronounced by native speakers of German or second language speakers with a Turkish accent, and then saw a face that was ethnically congruous or incongruous with the speaker’s accent. Their results showed that processing faces in incongruent pairs was cognitively more costly than in congruent pairs.

Stereotypes that influence the way we create expectations about what someone will say can therefore be activated extremely rapidly by their looks, but also by their accent. Furthermore, foreign or socially marked accents have been shown to be associated with negative stereotypes. Lev-Ari and Keysar (2010) have thus shown that trivial truisms (such as “ants do not sleep”) were more often evaluated as false when said with a foreign compared to a native accent (see however a response by Souza and Oliveira (2017) which suggests that this effect could be due to fluency and intelligibility matters rather than purely to the foreign accent). Dixon et al. (2002) and Dixon and Mahoney (2004) have also demonstrated that people speaking in a socially marked regional accent are more likely to be considered guilty of a violent crime than people speaking in a standard accent. People speaking slowly and with a standard accent are also most likely to be judged as trustworthy (Torre, 2017).

The present study aims at addressing two shortcomings of current research in the domain: the lack of studies on the activation of stereotypes by (1) the accent as opposed to the face, and (2) socially marked accents, as opposed to regional accents. In France, one of the most negatively-viewed accents is the so-called banlieue (suburban) accent, associated with young speakers from underprivileged urban areas around the largest cities – called *jeunes de banlieue* (literally, youth from the banlieue). It is characterised by pre-final lengthening (Fagyal, 2003), vowel reduction (Fagyal, 2010), palatalization/affrication of dental and velar plosives (Jamin et al., 2006), more constrictive and posterior realizations of /r/ (Jamin et al., 2006), and a prominent high-low pitch fall at major prosodic boundaries (Boula de Mareüil and Lehka-Lemarchand, 2011). In this stereotyped conception, the banlieue refers only to lower-class suburban areas with large housing estates and not to more middle-class or privileged suburbs which are also, in fact, banlieue towns. Young people from these underprivileged banlieues are constantly portrayed negatively in the media – they are relentlessly depicted as violent drug-addicts living in ghettos and falling prey to radicalized Islamism (Derville, 1997, 107). Derville also notes that the banlieue population is generally reduced to immigrants from Northern Africa, while foreigners only represent 23 % of the population in the Parisian Sensitive Urban Areas², and only around half of them are of Northern African origin (Chevalier and Lebeauvin, 2010). Longhi (2012, 142) also notes that the phrase *jeune de banlieue* conjures up very productive stereotypes relating to their being petty criminals or having no school leaving qualification. The banlieue only appears in news broadcasts in relation with violent events (Berthaut, 2013), which contributes to these stereotypes becoming entrenched. The accent of young people from working-class suburban areas is thus consistently associated with images of violence and crime.

The present study had two main objectives. We looked at the integration of accent-extracted information into the processing of the message, to contribute to the research that has been conducted on the unfolding of real-time language processing. We were particularly interested in the integration of social background information. Our second objective was to extend to the voice the research conducted on the activation of social and ethnic prejudice by faces. We were interested in the activation of negative stereotypes associated with criminality by the speaker's accent³. The study thus examined (1) the activation of social class stereotypes by the speaker's accent, with semantic violations triggered by the (in)congruence of the content with stereotypical expectations, and (2) the implicit activation of stereotypes associated with criminality by the banlieue accent.

For (1), we recorded EEG data from native speakers of French while they listened to self-referring sentences uttered by speakers with a banlieue accent or with an upper-class accent. These sentences could contain a mismatch between speaker accent and stereotypical content: a semantic content stereotypically associated with one of the groups (e.g., listening to rap or playing golf) was said in an accent associated with the other group. The critical comparison was between accent-content congruent and incongruent sentences. We looked at ERPs in the N400 (300-700 ms) and the P600 (700-900 ms) time-windows. Given previous research, and in particular the results of van Berkum et al (2008), we expected to observe an N400 incongruency effect after a mismatch between accent and semantic content (i.e., when self-referring stereotypical sentences consistent with the banlieue stereotype were uttered in an upper-class accent or when stereotypically upper-class content was

² *Zones Urbaines Sensibles* was the name given until 2014 to urban areas that were identified as facing important difficulties and therefore benefited from specific state help. They were characterized by old and degraded housing estates and a high rate of unemployment. These data from 2006 are the most recent ones that can be found on the origin of immigrants and foreigners in sensitive urban areas.

³ Stereotypes associated with criminality were chosen because this is a prominent feature of the stereotype associated with banlieue youth, as described above.

said in a banlieue accent). We did not expect to find a P600, as our anomalies were of a lexico-semantic and not of a syntactic nature, as pronouns would be. As a control, we also compared lexically congruent and incongruent sentences (e.g., “I eat pasta/*ashes almost every night”), which were used as distractors in the experiment. We expected to observe an N400 effect in response to these incongruencies.

For (2), we presented images after each sentence. These images either matched or did not match the content of the preceding stimulus. Half of the mismatching images were associated with criminal activities and the police (e.g., handcuffs, weapons, drugs) while the other half were unrelated to the content of the sentence and not evocative of criminal activities (examples are available in Figure 2). Participants were asked to judge whether the image matched the content of the sentence previously heard or not. We expected that if the banlieue accent was sufficient to trigger the activation of stereotypes associated with criminal activities, then participants would be slower to reject crime-related images after hearing a sentence in a banlieue rather than an upper-class accent. The critical comparison is therefore the effect of accent on the response time to crime-related images (note that these crime-related images were always incongruent with the content of the sentence, as no sentence contained a reference to criminal or illegal activities). We also expected a possible effect of accent on accuracy to these images: participants may be less accurate in rejecting these images, and accept more of them as congruent with the sentence content, when they followed a sentence in a banlieue accent.

2. Methods

2.1 Participants

26 native speakers of French (half male) took part in the EEG experiment. They were between 19 and 27 years old (mean 21;10) and reported being right-handed with normal hearing and vision. They all lived in the Paris area and attended university in Paris, except for one participant who lived in Strasbourg at the time of the experiment but spent most of her weekends in Paris. They grew up in different areas: 7 in Paris, 11 in the Greater Paris area, and 8 in other parts of France. To estimate participants' social background, they were asked about their parents' occupation. A standard 6-category frame was used⁴ (see Table 1): (1) Farmers, (2) Craftspeople, traders and business owners, (3) Senior executives and intellectual occupations, (4) Technicians, (5) Employees, and (6) Workers. Most participants reported that their parents belonged either to category 5 (48 %) or 3 (35 %). No participant spoke with a discernible banlieue accent⁵.

Participants gave written informed consent and were paid 20 euros. The procedure was approved by the local ethics committee (n°CER-PD:2019-14-AUTHORNAME). The data from two participants were excluded from further EEG analyses due to too many artefacts (>30%) in the EEG signal. The data

⁴ Categories from the INSEE classification were used (INSEE, 2003), although the last two (retired and no professional activity, which includes among other students) were not included. Participants were allowed to add information if they could not fit their case in the proposed categories. Although the wording of the categories is rather vague, participants are familiar with them from school. The INSEE category for Employees includes, among others, some categories of civil servants, police officers, or administrative employees in the private sector.

⁵ All participants were familiar with the banlieue accent, at least because some students at the university speak with this accent. However, their experience and familiarity with it may vary depending on their upbringing and the schools they attended. We unfortunately do not have detailed data on this point.

from 24 participants (12 male) were therefore analyzed for the EEG data, but the data from the other two participants were retained for behavioral analyses.

Socio-economic category	1	2	3	4	5	6
Parent 1	1	2	10	0	12	1
Parent 2	1	2	8	1	13	1
Total	2	4	18	1	25	2

Table 1 Participants' socio-economic background (1-Farmers, 2-Craftspeople, traders and business owners, 3-Senior executives and intellectual occupations, 4-Technicians, 5-Employees, 6-Workers)

2.2 Material

Stimuli

Following van Berkum et al. (2008), the target stimuli were 80 sentences including stereotypes matching a socially marked accent. Half were congruent with a Parisian upper-class accent (ex. “I love playing golf with my associates”), and the other half were congruent with the banlieue accent (“I always listen to rap music in my car”). Each sentence was recorded by someone whose accent matched the content and by someone whose accent did not match the semantic content. We avoided using marked slang words to maximize the chances that the incongruity would be semantic and not lexical in nature: it was the semantic content of the unexpected word that did not match the activated stereotype (ex. someone from a lower-class background is not expected to play golf), not the word itself⁶. Each sentence started with a neutral beginning of at least three words that could semantically match either accent, so that the listener would have time to extract accent information and implicitly activate stereotypical representations of the speaker. The duration of the initial segment did not differ between the Accent and Congruence conditions (Accent: $F(1, 156) = 1.02, p = .31$; Congruence: $F(1, 156) = 0.037, p = .85$; Accent:Congruence: $F(1, 156) = 0.28, p = .60$, see Table 2). The critical word was followed by a few words that ended the sentence to avoid the influence of sentence wrap-up processes, as in van Berkum et al (2008). Critical words of the two sets (banlieue/upper-class stereotypes) were matched for frequency^{7,8} ($t(51) = 0.48, p = .63$, estimated from Lexique 3.83 (New, 2006) from a corpus of film subtitles) and number of syllables ($t(77) = -0.89, p = .38$, see Table 2)⁹. Each participant heard only one version of each sentence, i.e., 40 socially congruent sentences, in which the content matched the social stereotypes associated with the speaker, and 40 incongruent ones.

	Banlieue set	Upper class set
Frequency (/million words)	36.64 (80.74)	23.20 (45.40)
Log frequency	2.13 (1.86)	1.91 (1.59)

⁶ As noted by one of the reviewers, the frequency of use of these particular words may still vary between the two groups of speakers, possibly making the incongruity lexical in some cases (e.g., a banlieue speaker is not expected to use the word “golf” itself). We did our best to use common words that could be used by both groups equally in different contexts, but the lexical aspect could have contributed to the effect in certain cases.

⁷ Note however that it was not possible to obtain frequency information for some of the critical items, namely proper nouns and acronyms (13 items for the banlieue set and 7 items for the upper class set).

⁸ Statistical analyses were carried out on the logarithmic transformation of the frequency.

⁹ Note that, due to the difficulty of constructing enough stimuli, we had to resort to using names of places, celebrities or brands, in particular for the Banlieue set (10 items, vs 7 items for the upper class set).

Number of syllables	1.97 (0.83)		2.15 (0.92)	
	Banlieue accent (Cong)	Upper-class accent (Inc)	Upper-class accent (Cong)	Banlieue accent (Inc)
Duration of the initial segment (ms)	853 (218)	842 (205)	804 (169)	827 (196)

Table 2 Frequency and number of syllables of critical stimuli and duration of the initial segment per accent set

96 socially neutral sentences containing a lexical-semantic incongruity as well as their 96 congruent counterparts were used as fillers. The incongruity was obtained by replacing one of the nouns or adjectives of the initial congruent sentence by another word which had the same number of syllables, a similar frequency of use (from Lexique 3.83; $t(189) = 0.04$, $p = .96$; $M_{\text{Congruent}} = 63.32$ per million words, $SD = 76.52$; $M_{\text{Incongruent}} = 62.82$, $SD = 81.19$) and the same grammatical category, but was semantically unacceptable (*Je suis allé à une **réunion**/***caméra** la semaine dernière.* – I went to a **meeting**/***camera** last week). These sentences were recorded by the same speakers as the target sentences, so as to reduce the saliency of the pairing between accents and stereotyped contents. Each pair was recorded by the same speaker, and each participant only heard one member of the pair. 48 congruent fillers and 48 incongruent fillers were therefore presented with the 80 target stimuli. A complete list of stimuli can be seen in Appendix A.

Voice selection

A total of 15 speakers with a perceived banlieue or upper-class accent were recorded in a familiar environment (their university or at home), so as to reduce the risk of the speaker adapting to the experimenter (this was a concern in particular for the banlieue accent). Each speaker recorded all the target stimuli twice and half of the fillers, for a total of 140¹⁰ sentences. The recordings were conducted with an AT2020 USB microphone and ROCme! Software (Ferragne et al., 2012). All recorded sentences were then cut to remove background noise, and their intensity was normalized to 70 dB with the Scale intensity function in Praat (Boersma and Weenink, 2016). The speakers who were recorded all gave written informed consent and were paid 5 euros for around 30 minutes of recording.

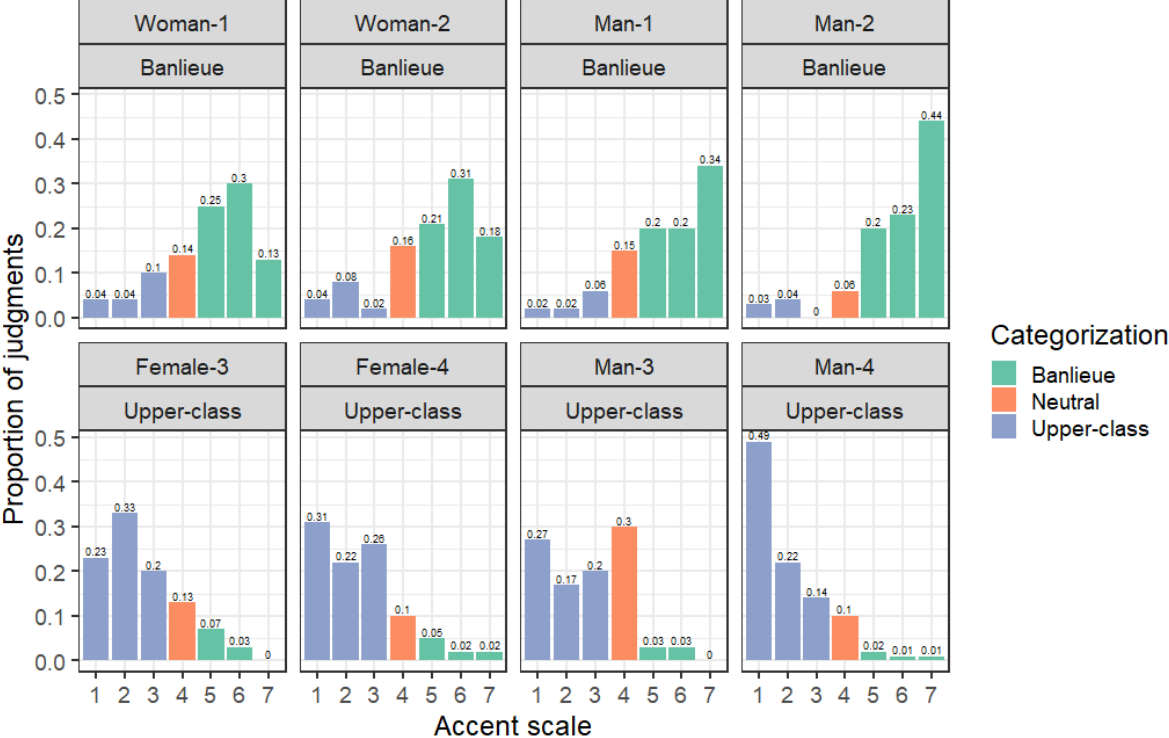
Although we only needed 4 speakers matching each social accent, we recorded 15¹¹ in total so that we could subsequently select the most stereotypical ones with the following procedure. Five recorded sentences were randomly chosen for each participant among the congruent fillers, so that the stereotypical content would not interfere with the perception of the accent. These sentences were presented auditorily to a naïve panel of 16 judges recruited among the colleagues and acquaintances¹² of the experimenter. Two drawings, representing a stereotypical youth from the banlieue and a stereotypical upper-class person, were presented on the left and right side of the screen and separated by 7 black squares. Judges were asked to rate how stereotypical the accent was, by clicking on a square towards the middle or closer to the corresponding picture depending on how close the accent of the speaker matched the stereotype corresponding to the picture. The speakers being most frequently classified at the ends of the stereotype continuum were selected.

¹⁰ Intentionally, more sentences were recorded than was needed (4 extra sentences per condition), so that we could eliminate potentially problematic sentences. This number corresponds to the recording of half of all critical stimuli (44×2) and a quarter of the fillers (52).

¹¹ Due to recruitment difficulties, only 3 men were recorded for the banlieue accent.

¹² Note that these judges were therefore mostly from middle to upper class backgrounds, as this was deemed to match the most frequent profile of the participants in the experiment. None of them spoke with a marked banlieue or upper-class accent – both accents were thus expected to be different from their own.

Analyses (a cumulative link mixed model with the `clmm` function from the `ordinal` package in R (Christensen, 2019), with a random intercept by Subject and Item and Intended accent as a fixed effect) confirmed that speakers were classified differently: the speakers chosen to represent the upper-class accent were rated 3.42 points lower (towards the upper-class end) on the Likert scale than the 4 selected for the banlieue accent (SE = 0.22, $z = -15.68$, $p < .0001$). The frequency of each rating for each speaker can be seen in Figure 1. We performed an auditory analysis of all the sentences, with particular attention to the production of /r/, potential affricate allophones of



plosives before /y/ and /i/ (vendredi: [vãdrødʒi]) with visual confirmation from spectrograms in Praat, and the presence or absence of optional liaisons. Liaisons are traditionally known to be optional between, for example, a verb form and a following word starting with a vowel: Je vais au lycée can either be realized as /ʒøvølise/ or /ʒøvøzlise/, the latter being favored by the higher social classes (Armstrong, 2001). Our auditory analyses confirmed that, in line with descriptions of the banlieue accent, the speakers produced typically salient realizations of /r/ and stereotypical affrication of plosives before /y/ and /i/. Banlieue speakers also produced fewer optional liaisons than upper-class speakers (banlieue speakers produced 2 (13 %) and upper-class speakers 12 (80 %) out of the 15 possible optional liaisons, see Table 3).

Figure 1 Frequency of ratings on the 7-point Likert scale for the perception of the accent of the 8 selected speakers. A rating of 1 corresponds to the most upper-class perception and a 7 corresponds to the most extreme banlieue perception. Giving the accent a rating between 1 and 3 was considered categorizing the speaker as upper-class, and between 5 and 7 correspond to a banlieue categorization. A rating of 4, in the middle of the scale, reflects a perception of the speaker as characteristic of neither stereotype¹³.

¹³ Note that Man 3 is classified as Neutral more often than other speakers. However, most of the judgements place him as an upper-class speaker, and not as a banlieue speaker. Due to the difficulty of finding available speakers matching our recruitment criteria, he was still included as an upper-class speaker. A model including Speaker as a fixed effect in addition to Stereotypical Congruence on the response time to the image association task revealed that there was no effect of the specific Speaker, either by itself ($\chi^2(7) = 8.99$, $p = .25$) or in interaction with Congruence ($\chi^2(9) = 9.96$, $p = .35$).

Speaker	W-1	W-2	M-1	M-2	W-3	W-4	M-3	M-4
Accent	B	B	B	B	UC	UC	UC	UC
Median rating	5	5	6	6	2	2	3	2
Number of optional liaisons produced	0	0	1	1	1	5	4	2

Table 3 Median rating and number of optional liaisons produced for each of the selected speakers. (B.: Banlieue accent, UC: Upper-class accent.)

Images

J'ai goûté un gâteau délicieux à midi.

*I tasted a delicious **cake** for lunch.*



a) Example of a congruent image

J'ai mis mon manteau en vison pour aller au théâtre

*I put on my **mink** coat to go to the theater.*



b) Example of an incongruent image

J'écoute toujours du rap dans ma voiture.

*I always listen to **rap** in my car.*



c) Examples of incongruent crime-related images

Figure 1 Examples of congruent, incongruent neutral and incongruent crime-related images used in the task

To check whether the banlieue stereotype as evoked by the banlieue accent would trigger the implicit activation of criminality-related stereotypes, like face-priming experiments have shown for racial prejudice, each sentence was also associated with a free stock image that could be either congruent or not with the semantic content of the stimulus (see Figure 2 for examples). Part of the incongruent images depicted situations or objects associated with a broad representation of

criminality (prison, handcuffs, arrest, etc.). No image depicted a crime that would be stereotypically associated with one or the other class (e.g., burning cars or tax evasion). None of the target stimuli contained any explicit reference to criminal activities, so any effect of speech on the reaction to the image could be attributed to stereotypes extracted from the accent and not directly from the semantic content.

Stimuli randomization

The final selection of critical stimuli contained an equal number of sentences for each of the 8 speakers (5 congruent and 5 incongruent stimuli by speaker). Each speaker also uttered an equal number of congruent and incongruent distractors (6 of each per speaker). The critical stimuli (40 congruent and 40 incongruent) as well as the distractor stimuli (48 congruent and 48 incongruent) were divided over two different lists containing an equal number of sentences from each condition and from each speaker. The lists were counterbalanced over participants. Each participant was presented with an individual pseudo-randomization of one of the lists. This pseudo-randomization was done with the following constraints: (1) no two items from the same condition¹⁴ could follow each other, (2) no more than two congruent or two incongruent stimuli followed each other (fillers and critical stimuli included) and (3) there were no more than two sentences uttered by the same speaker in a row. Although it was impossible to control for repetition on all dimensions of the experiment, we also tried to avoid large clusters of congruent or incongruent images (no more than 5 in a row), as well as the succession of crime-related images (presented on at most 3 successive trials for 3 participants, and at most 2 for the rest).

2.3 Procedure

Participants were seated in a soundproof booth in front of a computer. Stimuli were presented with Psychtoolbox-3 (Kleiner et al., 2007) on MATLAB R2016b (The Mathworks, 2016) through Sennheiser CX 1.00 in-ear headphones (28 Ω impedance and 119 dB). A fixation cross was displayed on the screen to reduce eye movements for 500 ms before the stimulus began, during stimulus presentation, and for 1000 ms after it ended. Then (1000 ms after the end of the stimulus), the corresponding image was displayed on the screen for 200 ms. Participants were then invited to judge whether it was congruent with the semantic content of the stimulus, and had to reply within 1000 ms by pressing the left or right arrow on the keyboard with, respectively, the index or medium finger of their right hand. The side corresponding to a “Yes” answer was counterbalanced across participants. Participants were then given the time they needed to blink.

EEG data recording and pre-processing

The EEG was recorded at a sampling rate of 1000 Hz at 24 bits with 64 electrodes placed according to the 10-20 standard, with a BrainVision BrainAmp amplifier and BrainVision pyCorder software, and processed with EEGLAB/ERPLAB (Delorme and Makeig, 2004; Lopez-Calderón and Luck, 2014). The electro-oculogram (EOG) was obtained from an electrode placed under the left eye. Impedance was kept under 16 k Ω . Epochs from -300 to 1200 ms around the critical point were extracted from continuous data and linearly detrended with the EEG linear detrend function from ERPLAB. The data were baseline-corrected (-300 – 0 ms), bandpass filtered offline at 0.1-30 Hz and re-referenced to the average of both mastoids. Artefact detection was performed with the following procedure. First, the

¹⁴ The conditions were defined as the following: (1) Upper class set, congruent accent; (2) Upper class set, incongruent accent; (3) Banlieue set, congruent accent; (4) Banlieue set, incongruent accent; (5) Congruent distractor; (6) Incongruent distractor.

blink rejection algorithm from ERPLab was used to detect 100 ms blinks. A 200 ms moving window with a 50 ms step was then run on all electrodes to detect other artefacts. The threshold was adapted to participants to maximize the detection of artefacts and varied between 80 and 100 μ V.

To correct for latency variability in ERP components, we used the Residue Iteration Decomposition (RIDE) method developed by Ouyang, Zhou and Sommer (Ouyang et al., 2011a, 2013, 2015, 2016). This method separates clusters of superimposed components that could be related to different cognitive processes. Ouyang et al. (2016) show that decomposing and then reconstructing ERPs with RIDE can restore condition effects when they are blurred by variability in component latencies, without distorting the rest of the signal. RIDE was applied to the data with the RIDE toolbox for MATLAB (Ouyang et al., 2011b) separately for each condition (Semantic Congruent / Semantic Incongruent / Social Congruent / Social Incongruent), and was set to search for one component in a 200-800 ms time-window.

ERPs were time-locked to the beginning of the critical word. Following van Berkum et al. (2008), we selected a 300-700 ms time-window for a large N400 effect. We additionally investigated possible P600 effects in a 700-900 ms window. The critical results are therefore whether the average amplitude in the given time-window differs significantly between the congruent and incongruent conditions, and whether this difference is present in all the determined surface regions. The dependent variable was therefore the average amplitude across trials for each of the critical conditions separately (congruent vs. incongruent) in the given time-window¹⁵, as those two congruence conditions were compared. Lateral and medial electrodes were examined in two separate analyses to look at possible differences between hemispheres. Medial electrodes were grouped into 3 regions of interest (ROIs; see Table 4): Anterior, Central, and Posterior. Lateral electrodes were divided into 6 standard regions of interest, allowing us to look at effects of anteriority and hemisphere on the potential N400 effect (see Table 4)¹⁶. Analyses were conducted on the average amplitude in the given time-window with linear mixed-effect models in R with the lmer() function from the lme4 package version 1.1-21 (Bates et al., 2015) with Congruence (Congruent / Incongruent) and Region (Anterior Left / Anterior Right / Central Left / Central Right / Posterior Left / Posterior Right for the lateral analysis; or Anterior / Central / Posterior for the midline analysis) and their interaction as fixed effects, and a random intercept per participant as well as random slopes whenever possible for Congruence and Region, which were both within-subject factors¹⁷ (see complete models in Appendix B). Separate models were run for Stereotypical incongruence and Semantic incongruence. All contrasts were coded as sum contrasts. The significance of main effects and interactions was evaluated through log-likelihood model comparison with the R function anova(). As a main effect of the Region is not of interest to our question (it would only mean that the amplitude is higher in some areas, because of general language treatment or brain activity unrelated to the task), it is not discussed further. Follow-up analyses of significant Congruence:Region interactions were conducted with nested contrasts models making use of contrast coding instead of running several submodels (Schad et al., 2020), as we were only interested in the effect of Congruence for each level of the Region variable (i.e., whether there was a different electrical activity

¹⁵ Please note that this means there is one measure for each electrode and each condition per subject. Single-trial analyses are usually not done in ERP analyses because averaging across trials allows us to cancel out non-experiment related background noise (Luck, 2014, 23). This means that there is no random effect by item.

¹⁶ Although electrodes are grouped by region of interest, measurements are not averaged for each region but simply coded for ROI, which means that we retain one observation per electrode.

¹⁷ Due to convergence and singularity issues, it was not possible to add a random slope for all the terms that would ideally be included in a maximal random-effects structure (a slope for Congruence, a slope for Region, and a slope for their interaction).

in response to matching and mismatching stimuli in the different ROIs, and not whether the activity differed in the various ROIs for each level of Congruence separately).

	Left	Midline	Right
Anterior	AF3, AF7, F1, F3, F5, F7	AFz, Fz	AF4, AF8, F2, F4, F6, F8
Central	FC1, FC3, FC5, C1, C3, C5, CP1, CP3, CP5	FCz, Cz, CPz	FC2, FC4, FC6, C2, C4, C6, CP2, CP4, CP6
Posterior	P1, P3, P5, P7, PO3, PO7	Pz, POz	P2, P4, P6, P8, PO4, PO8

Table 4 Grouping of electrodes by Region of Interest

3. Results

3.1 Behavioral task

Response times inferior to 150 ms and for incorrect answers were removed. Accuracy data when the response time was inferior to 150 ms was also removed. Participants were attentive and performed well on the image congruence task ($M = 89.58\%$, $SD = 5.41$, range = 76.14 – 97.16 %).

Analyses were conducted in R with linear mixed-effect models, with the lme4 package. Models were run on the logarithmic transformation of the response time to achieve normality of the residuals. Accuracy was examined by means of a binomial model with a logit link with the glmer() function. The maximal random effects structure that converged and was not singular was included (details of the random effects structure for each model are provided in Appendix B, Tables 1.1 to 1.3).

We first looked at whether the response time to the images was affected by the mismatches in the sentences. If the mismatches were disruptive enough, they should slow down participant's performance— we would thus expect sentences containing an incongruence to be followed by longer response times on the image association task, regardless of whether the image matched the content or not. We looked separately at the stereotypical incongruences (between stereotypes and accent) and the semantic incongruences. Response times and accuracy data per congruence and type of mismatch are reported in Table 5. There was no effect of Stereotypical Congruence nor Accent on the reaction time (all $ps > .26$)¹⁸, nor on the accuracy (all $ps > .39$). The response time was however affected by Congruence for the lexico-semantic items ($\chi^2(1) = 12.82$, $p = .0003$): it took participants on average 32 ms longer to react to the image following a semantic incongruence (e.g., "I went to a ***camera** last week"; $\beta = 0.034$, $SE = 0.009$, $M_{\text{Congruent}} = 370 (\pm 149)$ ms, $M_{\text{Incongruent}} = 402 (\pm 172)$ ms). The effect of Congruence was not significant on Accuracy performance ($\chi^2(1) = 2.64$, $p = .10$). The semantic mismatches were severe enough to interfere with the following task, even though the image only appeared 1 s after the end of the stimulus. The stereotypical mismatches were, however, not strong enough to have the same effect.

	Semantic mismatches		Stereotypical mismatches	
	RT (ms)	Accuracy (%)	RT (ms)	Accuracy (%)
Congruent content	370 (149)	98.15 (2.20)	364 (151)	94.00 (6.87)
Incongruent content	402 (172)	89.58 (9.47)	366 (155)	94.33 (5.30)
Difference	+32 **	-8.57	+2	+0.33

¹⁸ The model included a fixed effect for Congruence and for Accent as well as their interaction, to control for the eventuality that stereotypical mismatches were stronger for one accent than for the other.

Table 5 Response time and performance per congruence of the content and type of mismatch (mean and standard deviation)

We then looked at the data for crime-related images only, as our hypothesis was that participants would be slower to reject these images after hearing a Banlieue rather than an Upper-class accent, if they implicitly associated the stereotype associated with the Banlieue accent with a representation of criminal activities. Response times and accuracy levels by image congruence and speaker accent are reported in Table 6. Analyses were run on the full dataset, including both stereotypical and semantic items, as the effect of mismatch was not included in the models, and as there were not enough crime-related images for each condition to be examined separately. A model run on the logarithmic transformation of the reaction time with Accent as a fixed effect did not reveal a significant effect of Accent on the speed of rejection of crime-related images ($p = .38$). The binomial model on the accuracy data did not reveal any effect of Accent either ($p = .26$).

Accent	RT (ms)		Accuracy (%)	
	Banlieue	Upper-class	Banlieue	Upper-class
Congruent image	372 (163)	375 (163)	92.61 (4.69)	94.32 (8.04)
Incongruent – neutral	385 (153)	372 (149)	98.05 (2.43)	95.91 (6.93)
Incongruent – crime	371 (150)	385 (160)	97.73 (2.43)	97.16 (3.38)

Table 6 Response time and performance per congruence of the image and accent (mean and standard deviation)

3.2 EEG

Semantic incongruencies

300-700 ms Inspection of the waveforms suggests the presence of the expected N400 effect in the central region (see blue rectangles on *Figure 3*)¹⁹. Analyses confirmed that the Congruence:Region interaction was significant in the 300-700 ms time window over lateral electrodes ($\chi^2(5)=16.75, p = .005$, see full results in Appendix B, Table 2.1)²⁰: there was a significant difference between the Incongruent and Congruent conditions in some but not all of the Regions of interest. Semantic incongruencies (“I broke a *thought/bottle of champagne”) triggered a significant N400 effect (i.e., amplitude in the Incongruent condition was more negative than in the Congruent condition) in the central region, both left ($M_{\text{Incongruent-Congruent}} = 0.88 \mu\text{V}, SE = 0.32, t(27) = 2.24, p = .01$) and right ($M_{\text{Inc-Cong}} = -0.72 \mu\text{V}, SE = 0.32, t(27) = -2.24, p = .03$). There was a similar trend in the posterior left region ($M_{\text{Inc-Cong}} = -0.57 \mu\text{V}, SE = 0.33, t(29) = -1.74, p = .09$). The Congruence:Region interaction was also significant at midline sites ($\chi^2(2) = 7.10, p = .029$): there was a significant N400 effect in the anterior ($M_{\text{Inc-Cong}} = -1.13 \mu\text{V}, SE = 0.43, t(27) = -2.61, p = .015$) and central regions ($M_{\text{Inc-Cong}} = -0.98 \mu\text{V}, SE = 0.43, t(26) = -2.30, p = .030$).

700-900 ms Analyses in the 700-900 ms window revealed a significant Congruence:Region interaction at lateral ($\chi^2(5) = 159.87, p < .0001$) and medial electrodes ($\chi^2(2) = 31.99, p < .0001$). At lateral sites,

¹⁹ Note that what matters is the difference between the Congruent and Incongruent waveforms. This means that an N400 effect will be reflected by Incongruent values (red solid curve) that are more negative (i.e., lie above, as the negative values are plotted up) than Congruent values (black dotted curve).

²⁰ As visual inspection of the waveforms suggested that the N400 effect might be preceded by a small negativity at anterior sites, we ran an additional model in a 0-300 ms window, including Congruence, Region and their interaction as fixed effects, which revealed a significant Congruence:Region interaction at lateral ($\chi^2(5) = 58.22, p < .0001$) and medial sites ($\chi^2(2) = 27.11, p < .0001$). This additional negativity was only significant in the anterior medial and right ROIs (see Appendix B, Table 2.3 for full details).

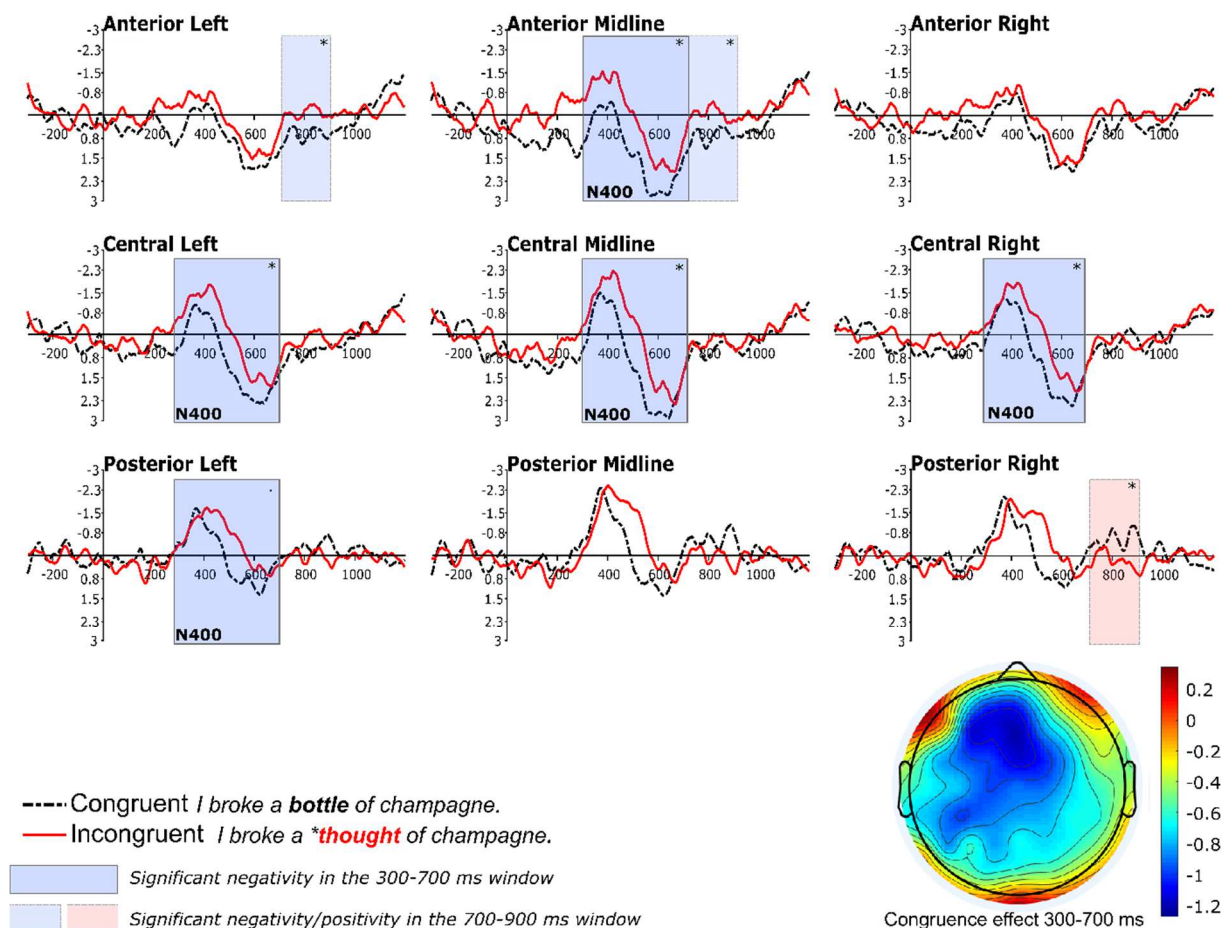


Figure 2 ERP waveforms and scalp map for semantic incongruities (Horizontal axis: time from -300 to 1200 ms; Vertical axis: amplitude from -3 to 3 μV . Negative values are plotted up. Scalp map: difference between Incongruent and Congruent conditions between 300 and 700 ms. The terms “positivity” and “negativity” refer to the difference between the average amplitude in the Incongruent and Congruent conditions in the given time window, not to the absolute values.)

the Congruence effect was only significant in the anterior left area, revealing a more negative amplitude in the Incongruent condition ($M_{Inc-Cong} = -0.76 \mu\text{V}$, $SE = 0.31$, $t(29) = -2.46$, $p = .020$, see Appendix B, Table 2.2, and light blue and red rectangles on Figure 3). In the medial area, the Congruence effect revealed a negative deflection in the Incongruent condition in the anterior region ($M_{Inc-Cong} = -0.78$, $SE = 0.37$, $t(33) = -2.11$, $p = .043$), but a slight positive one in the posterior right area ($M_{Inc-Cong} = 0.72 \mu\text{V}$, $SE = 0.31$, $t(29) = 2.32$, $p = .028$).

Stereotypical incongruities

Inspection of the waveforms suggested the co-occurrence of two components: an N400-like antero-central negativity, mostly localized in the left hemisphere (i.e., a more negative response in the Incongruent than the Congruent condition, see blue rectangles on Figure 4)²¹, and a simultaneous P3 (i.e., a positive difference due to a more positive response in the Incongruent than the Congruent condition, see red rectangles on Figure 4)²².

²¹ Note again that what matters is the difference between the two conditions, and whether the amplitude in the Incongruent condition (solid red curve) is more negative than in the Congruent condition (dotted black curve).

²² As suggested by one of the anonymous reviewers, we conducted supplementary analyses to check for a possible influence of the socio-economic background of the participants on the congruence effect. A socio-economic variable was included as a fixed effect in the models. This variable was a simple binary factor

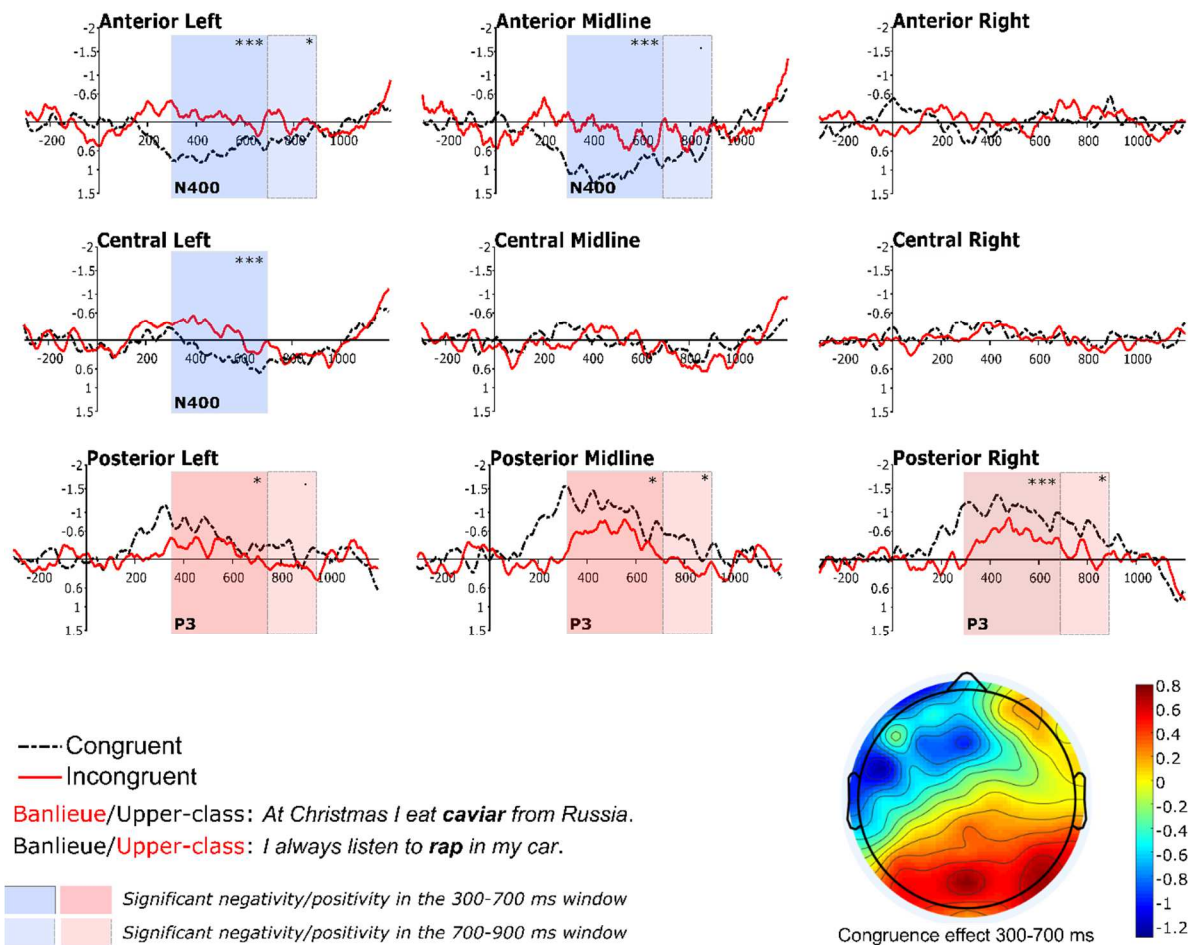


Figure 3 ERP waveforms and scalp map for stereotypical incongruities (Horizontal axis: time from -300 to 1200 ms; Vertical axis: amplitude from -2 to 1.5 μV . Negative values are plotted up. Scalp map: difference between Incongruent and Congruent conditions between 300 and 700 ms. The terms “positivity” and “negativity” refer to the difference between the average amplitude in the Incongruent and Congruent conditions in the given time window, not to the absolute values.)

300-700 ms Statistical analyses showed a significant Congruence:Region interaction in the a priori designated time-window (300-700 ms) at lateral ($\chi^2(5) = 64.74, p < .0001$) and midline sites ($\chi^2(2) = 24.49, p < .0001$, see full results in Appendix B, Table 3.1)²³. The Congruence effect was significant in most Regions of interest, but going in two different directions. Social incongruities (e.g. “At Christmas I eat caviar from Russia” in a banlieue accent) triggered a negative difference effect in the anterior left ($M_{\text{Inc-Cong}} = -0.70 \mu\text{V}, SE = 0.16, t(199) = -4.50, p < .001$) and central left regions ($M_{\text{Inc-Cong}} = -0.48 \mu\text{V}, SE = 0.13, t(113) = -3.59, p < .001$), as well as in the anterior midline area ($M_{\text{Inc-Cong}} = -0.81 \mu\text{V}, SE = 0.23, t(90) = -3.56, p < .001$). At posterior sites however, those incongruities were followed by a positive difference (i.e., values in the Incongruent condition were higher than in the Congruent

classifying the participants as belonging to the upper class or not. Participants who had both parents in category 3 (senior executives and intellectual occupations) were classified as having an upper-class background. These supplementary analyses (reported in Appendix B, section 4) did not reveal any effect of socio-economic background.

²³ As visual inspection of the waveforms suggested that the components may start around 100 ms after the critical point, we conducted supplementary analyses in the 100-300 ms time window. These revealed a significant Congruence:Region interaction ($\chi^2(2) = 25.67, p < .0001$, see Appendix B, Table 3.3 for full details): the positivity started early in the posterior Left, Midline and Right ROIs (all $ps < .05$), and the negativity started early in the anterior left and midline ROIs ($p < .05$).

condition). This positive difference was significant in the posterior right ($M_{\text{Inc-Cong}} = 0.56 \mu\text{V}$, $SE = 0.16$, $t(199) = 3.57$, $p < .001$), medial ($M_{\text{Inc-Cong}} = 0.55 \mu\text{V}$, $SE = 0.23$, $t(90) = 2.44$, $p = .02$), and left ($M_{\text{Inc-Cong}} = 0.35 \mu\text{V}$, $SE = 0.16$, $t(199) = 2.26$, $p = .025$) regions. This posterior positivity was reminiscent of a P3, a component usually associated with context-updating and processing of task-relevant stimuli (Donchin, 1981).

700-900 ms Analyses in the later time-window (700-900 ms, see light blue and red rectangles on Figure 4 and Appendix B, Table 3.2) showed a significant Congruence:Region interaction for lateral ($\chi^2(5) = 42.80$, $p < .0001$) and medial electrodes ($\chi^2(2) = 16.35$, $p = .0003$). There was a significant Congruence effect in the anterior left region, revealing a prolonged negativity in that area ($M_{\text{Inc-Cong}} = -0.32 \mu\text{V}$, $SE = 0.15$, $t(70) = -2.10$, $p = .040$). At medial sites, the effect of Congruence was only marginally significant in the anterior region ($M_{\text{Inc-Cong}} = -0.37 \mu\text{V}$, $SE = 0.22$, $t(58) = -1.70$, $p = .095$, see Appendix B, Table 3.3). The earlier positivity was sustained in that window at posterior sites (Left: $M_{\text{Inc-Cong}} = 0.30 \mu\text{V}$, $SE = 0.15$, $t(70) = 1.95$, $p = .055$; Medial: $M_{\text{Inc-Cong}} = 0.50 \mu\text{V}$, $SE = 0.22$, $t(58) = 2.33$, $p = .023$; Right: $M_{\text{Inc-Cong}} = 0.52 \mu\text{V}$, $SE = 0.15$, $t(70) = 3.38$, $p = .001$).

The stereotypical mismatches thus triggered two separate components. The expected N400 had a non -typical distribution, being mostly left-distributed, and was more sustained than for semantic violations, starting around 100 ms, and continuing until around 900 ms after target-word onset at frontal sites. The P3 appearing at posterior sites around 100 ms after the critical point was maximal around 300 ms and also sustained until 900 ms.

4. Discussion

In this study, we investigated the way social stereotypes are processed in real time and can influence the interpretation of a semantic content. To this end, participants coming mostly from a middle to upper-class background²⁴ listened to sentences where the semantic content either matched the stereotypes associated with the accent of the speaker (ex: “I eat caviar from Russia” said in an upper-class accent, or “I always listen to rap in my car” in a lower-class, negatively connoted banlieue accent) or didn’t match those stereotypes (i.e., “I eat caviar from Russia” said in the banlieue accent). We analyzed ERPs triggered by those incongruencies and by classic semantic mismatches (“I went to a meeting/*camera last week”). We also looked at accuracy and response time in rejecting crime-related images incongruent with the stimulus heard.

Behavioral results did not show an effect of accent on the processing of crime-related images. We expected crime-related images to be rejected faster after a stimulus said in an upper-class accent than a banlieue accent, like tools are correctly identified (as tools rather than weapons) faster after seeing the face of a White rather than Black person. There are several possible reasons for these non-results that make them difficult to interpret. First of all, unlike in classic behavioral experiments, a blank of 1000 ms separated the end of the sentence and the presentation of the image in our EEG task, so that response processes would not contaminate reactions to the critical stimuli. This probably gave participants enough time to consciously process the sentence they had heard and to repress implicit prejudice. The variety of images used may also have caused a low signal-to-noise ratio, as some images may have been more easily associated with criminal activities or with one or the other social groups, and so not necessarily triggering the expected stereotype. In the future, thorough ratings should be obtained for the images to determine how strongly they evoke criminal activities and certain social backgrounds. In addition, as we wanted to include the

²⁴ Note however that one participant was from a working-class background.

image task as a distractor task in the EEG experiment, it was dependent on the stimuli used for the latter. This means that crime-related images could appear after semantically or stereotypically congruent and incongruent content, which introduces variability. It would be worthwhile in future research to conduct this image task separately from EEG data collection, with better controlled images and with only stereotypically neutral and semantic congruent sentences.

Semantic mismatches triggered an N400 effect, maximal at central electrodes. Additional analyses showed that this N400 effect was preceded by a small negativity in the anterior medial and right regions – this negativity was sustained in the N400 window in the medial area, but not in the right hemisphere. It is difficult to explain why any response to the semantic mismatch would start so early – immediately at the beginning of the critical word. A first explanation is that it could be an artefact due to different reactions to the previous word. However, as we manipulated the target and not the context, the word preceding the congruent and incongruent version of each item was the same, and so cannot have triggered different N400 effects – a problem that has been invoked to explain some early negativities (Steinhauer and Drury, 2012). There could still be some slight prosodic differences between the two conditions, causing the listeners to anticipate the incongruent content. There is, however, an alternative explanation. Frontal and in particular left frontal sustained negativities have been linked to increased working-memory demands in the retention interval. For example, Ruchkin et al. (1992) found that having to retain longer non-words led to an increased anterior negativity in the retention interval – which they associate with phonological loop rehearsal. Fiebach et al. (2001) and Phillips et al. (2005) have observed sustained anterior negativities following the initiation of wh-dependencies, negativities which are larger when there is more demand on working memory (i.e., in the case of object- vs subject-relatives). It is possible that we are observing two parallel components: (1) a centrally-distributed N400 effect in response to the lexico-semantic violations, significant in the traditional time-window, and extending to posterior areas, and (2) an anterior negativity, starting early and sustained until around 900ms at medial sites, reflecting memory processes. This might be due to specific task-demands: as the image association task required keeping the sentence in working memory to judge the congruence of the following picture, incongruent words might have made the content harder to memorize – and, possibly, been anticipated as potential keywords to be matched with the subsequent image. This is in line with the behavioral results: participants were slower to react to the image after semantically incongruent content, possibly because incongruencies made the content harder to encode and to retrieve.

We found that listeners were sensitive to social incongruencies between the accent of the speaker and the stereotypical semantic content of sentences. Mismatches triggered a posterior positivity and a negativity that was more anterior and more left-lateralized than the N400 effect that was expected given the results obtained by van Berkum et al. (2008). There are two possible interpretations of this distribution. A first possibility is that this negative shift is not an N400 effect, but a left-lateralized sustained negativity reflecting increased working-memory demands (Ruchkin et al., 1992; Fiebach et al., 2001; Phillips et al., 2005). This could be related to specific task demands – participants may have found it harder to remember stereotypically incongruent content to judge whether it matched the following picture. However, they were not slower to judge images after stereotypically incongruent content. Another possible explanation is that stereotypical mismatches did trigger an N400 effect, but that its presence at centro-posterior sites is hidden by the presence of the concurrent P3. The fronto-central distribution may also stem from the fact that auditory N400s also tend to be more frontal and less right-biased than visual N400s (Kutas and Federmeier, 2011). Fronto-central N400s have also been found for stereotype-related mismatches. Proverbio et al. (2017) visually presented sentences whose last word either fulfilled or contradicted gender-based occupational stereotypes (e.g., “The notary is breastfeeding”). They observed that stereotype

violations triggered an anterior N400 effect maximal between 350 and 450 ms after the incongruency. Hansen et al. (2017) found a fronto-central N400 effect when the face of the speaker did not match the ethnic expectations associated with their German or Turkish accent. The N400 effect obtained here suggests that participants are considering the critical word as implausible – or at least less predictable – when associated with the stereotypically incongruent accent, which entails that they had created expectations about the upcoming word based on their representation of the speaker.

The accent-content mismatches also triggered a salient posterior positivity maximal between 200 and 400 ms after the onset of the target word, suggesting a P3. This component is thought to reflect the updating of context-related expectations (Donchin, 1981), and it has also been found in the case of stereotypical mismatches. Bartholow et al. (2006) conducted a face priming task designed to elicit stereotype activation and found that counter-stereotypical associations triggered larger and longer P3s. Jia et al. (2012) also observed a P3 when non-stereotypic traits for Tibetans followed stereotypical Tibetan names (when judged by Han Chinese participants). The P3 observed in response to stereotypical mismatches has been claimed to reflect semantic knowledge updating. Ito and Bartholow (2009) argue that stereotype incongruencies challenge our existing semantic knowledge, and therefore engage mechanisms – in particular the locus coeruleus-norepinephrine system, which has been linked to P3 activity – to update the existing semantic content or attempt to resolve the inconsistent information. Mitchell et al. (2009) also observed that stereotype application (using stereotypes to infer someone's preferences and opinions) activated a large area in the right frontal cortex, which has been linked among other things to semantic retrieval of categorical knowledge (e.g., which social category the speaker belongs to). Our data suggests that a P3 can also be obtained from stereotypes carried by accent, not just faces or family names. The semantic mismatch between the content of the message and the stereotypes associated with the social group of the speaker causes a need to check and update semantic knowledge about this group. The presence of a P3 in our data but not in van Berkum et al's study might be due to the presence of the image association task. Participants in van Berkum et al (2008) only had to passively listen to the sentences, while ours were required to make an – apparently – unrelated judgment. However, the presence of this task may have triggered stereotype application (i.e., using semantic knowledge about the represented speaker associated with the accent to evaluate whether the image matched the content, even though there was actually no link between the accent and the congruence of the image) and not just activation (i.e., unconsciously associating the representation of crime with the stereotype associated with the banlieue accent), as the crime-related pictures may have prompted participants to think about group prejudice.

It is difficult to directly compare the semantic and social N400 effects obtained in our experiment, as we did not plan on comparing effect sizes, and the semantic mismatches were more unacceptable than the social ones, leading to larger effects. The concurrent presence of the P3 also makes it difficult to quantitatively compare the two effects. Like van Berkum et al. (2008), we scaled the data by dividing each electrode-specific value of the difference effect for each participant and type of incongruency by the mean for that participant and that incongruency type, to adjust for the difference size in the two effects. Unlike van Berkum et al., we still found a difference in the scalp distribution of the two effects (Incongruency type:Region, lateral electrodes: $\chi^2(10) = 22.46, p = .013$; medial electrodes: $\chi^2(4) = 4.48, p = .34$). This difference was limited to the anterior left area ($p < .01$), where only the social N400 was significant. It is possible that this reflects the fact that processing the social mismatches leads to additional mechanisms being recruited compared to a typical semantic N400; however, the concurrent presence of the P3 makes it difficult to draw strong conclusions

based on topography. A replication of the study would be needed to confirm the localization of this social N400.

Our results obtained with incongruencies based only on social stereotypes are generally consistent with those of the study van Berkum et al. (2008) conducted with a mix of social, gender and age mismatches. This confirms that after only a few words and thus a very short amount of time (829 ± 202 ms on average here), listeners have already formed expectations about what the speaker is likely to say, based on their accent alone and the representations evoked by that accent. The N400 is commonly associated with the difficulty of accessing and integrating conceptual information stored in semantic memory (Kutas et al., 2006), which includes stereotypes about different social groups, and to the predictability of the upcoming word. It is therefore not surprising that the N400 effect has now repeatedly been associated with stereotype accessibility in paradigms pairing words (White et al., 2009; Wang et al., 2017) or words and faces (Hegeman et al., 2014).

As in Hegeman et al.'s study, the association of accent and stereotypes was socio-politically charged, and it is important to note that the ERP responses appear even though participants might be trying to tune down their stereotypes. Our experiment shows that a few seconds of speech – without any socially marked vocabulary, and without any hint as to the visual appearance of the speaker – are enough to trigger socially-based predictions as to the content of the speaker's message. If predictions can of course be overcome, recent studies show that expectations about the speaker can persist and lead to discrimination. Of crucial importance, several studies have found that dialect and prejudice associated with it can influence juror decisions in criminal trials: Kurinec and Weaver (2019) thus showed that, in a mock ambiguous criminal case, defendants speaking African American Vernacular English were evaluated more negatively and found guilty more often than speakers of General American English. Similarly, Cantone et al. (2019) found that Mexican Americans and Black Americans were judged more negatively than White Americans especially when Black Americans spoke with a stereotypical accent. These studies demonstrate that dialect and accent-related prejudice can have dire consequences for their speaker in particular in the judicial system, and underlie the need for more research and to raise awareness of these implicit biases. It is worth noting here that in our study, although we did our best to ensure that the incongruity would be based on semantic expectations (i.e., specific activities stereotypically associated with each social background), we cannot entirely exclude that it may have been partially lexical in some cases. Indeed, in the absence of frequency data on the use of the critical items by members of each social class, we cannot guarantee that what was unexpected by the listener was the semantic content (e.g., a banlieue youth eating caviar) and not the use of the word itself (e.g., a banlieue youth is not expected to use the word 'caviar'). As the N400 effect has been observed for lexical as well as semantic effects, this does not change the nature of the mechanisms observed – the conclusion that listeners form expectations about the content of the speaker's message early and based on their accent alone still holds. However, the nature of the stereotypes would arguably be different – the expectations formed by the listener would then be lexical in nature, and linked to the form of the message, rather than semantic and relating to expectations about activities. It is also worth reminding the reader here that the stereotypes used in the EEG stimuli did not refer to one of the most prevalent negative stereotypes associated with banlieue youth, i.e., criminality. The stereotypical mismatches reflected by the EEG data are barely associated with socio-economic background, and we have no data supporting the association of implicit stereotypes related to criminal activities with the banlieue accent itself. Further research will be needed to (1) establish whether the stereotypical incongruity effect obtained here is lexical or semantic in nature and (2) the nature of the stereotypes implicitly triggered by the banlieue accent of the speaker. Note however that studies with stimuli similar to ours on principle, i.e., where the mismatch was based on

a discrepancy between the voice of the speaker and the content of a self-referring statement (Lattner & Friederici, 2003; Van Berkum et al., 2008), considered the effect as being semantic and not lexical.

ERPs can provide an effective tool to assess such social biases and to determine what exactly triggers them. Hehman et al. (2014) note that the N400 can function as an index of stereotype accessibility in the interracial domain, but we show that this can be extended to the social domain – even though there is definitely an ethnic prejudice component in the stereotypes associated with the banlieue accent, as Île-de-France suburbs, which have the highest migrant populations in the country (Armstrong and Pooley, 2010), are commonly reduced to their Northern African inhabitants (Derville, 1997). Comparing purely social dialects without the racial undertone might be of interest to further our understanding in this domain. We found that the P3 component can also serve as an indicator of stereotype knowledge retrieval. Hehman et al. have observed that N400 activity is moderated by self-reported explicit racist bias. We unfortunately did not collect explicit prejudice bias in this study, but this might prove useful information and would be worth collecting in further studies.

5. Conclusion

Our study shows that socially marked accents can activate expectations about the content of the speaker's message after an extremely short time. Our results are in line with other studies suggesting that stereotype knowledge is part of semantic categorical knowledge and that stereotypical mismatches may trigger the retrieval of this knowledge for checking and updating purposes. Prejudice about accents and dialects can have extremely serious consequences, in particular in the courtroom, and professionals need to be made aware of how hearing the defendant speak may influence their decision.

Acknowledgements

This work was supported by the Agence Nationale de la Recherche VOXCRIM project (ANR-17-CE39-0016). We are grateful to our two anonymous reviewers for their comments and feedback.

References

- Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience* 15, 670–682. doi:10.1038/nrn3800.
- Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., and Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science* 15, 88–93. doi:10.1111/j.0963-7214.2004.01502003.x.
- Armstrong, N. (2001). *Social and Stylistic Variation in Spoken French*. Amsterdam: John Benjamins Publishing Company doi:10.1075/impact.8.
- Armstrong, N., and Pooley, T. (2010). *Social and Linguistic Change in European French*. Palgrave Macmillan UK doi:10.1057/9780230281714_1.

- Bartholow, B. D., Dickter, C. L., and Sestir, M. A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology* 90, 272–287. doi:10.1037/0022-3514.90.2.272.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., et al. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package, version 1.1-9. Available at: <https://cran.r-project.org/web/packages/lme4/index.html>.
- Berthaut, J. (2013). *La banlieue du "20 heures" : ethnographie de la production d'un lieu commun journalistique*. Agone. Available at: <https://hal.archives-ouvertes.fr/hal-01138646/> [Accessed January 24, 2019].
- Boersma, P., and Weenink, D. (2016). Praat: doing phonetics by computer. Version 6.0.22.
- Boula de Mareüil, P., and Lehka-Lemarchand, I. (2011). Can a prosodic pattern induce/reduce the perception of a lower-class suburban accent in French? *Proceedings of the 17th ICPHS*, 348–351.
- Canal, P., Garnham, A., and Oakhill, J. (2015). Beyond gender stereotypes in language comprehension: Self sex-role descriptions affect the brain's potentials associated with agreement processing. *Frontiers in Psychology* 6, 1953. doi:10.3389/fpsyg.2015.01953.
- Cantone, J. A., Martinez, L. N., Willis-Esqueda, C., and Miller, T. (2019). Sounding guilty: How accent bias affects juror judgments of culpability. *Journal of Ethnicity in Criminal Justice* 17, 228–253. doi:10.1080/15377938.2019.1623963.
- Chevalier, C., and Lebeaupin, F. (2010). La population des zones urbaines sensibles. *INSEE Premiere* 1328. Available at: <https://www.insee.fr/fr/statistiques/1280948>.
- Christensen, R. H. B. (2019). ordinal - Regression models for ordinal data. R package version 19.12-10. Available at: <https://cran.r-project.org/package=ordinal>.
- Clopper, C. G., and Bradlow, A. R. (2009). Free classification of American English dialects by native and non-native listeners. *Journal of Phonetics* 37, 436–451. doi:10.1016/j.wocn.2009.07.004.
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods* 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009.
- Derville, G. (1997). La stigmatisation des « jeunes de banlieue ». *Communication et langages* 113, 104–117. doi:10.3406/colan.1997.2789.
- Dixon, J. A., and Mahoney, B. (2004). The effect of accent evaluation and evidence on a suspect's perceived guilt and criminality. *Journal of Social Psychology* 144, 63–73. doi:10.3200/SOCP.144.1.63-73.
- Dixon, J. A., Mahoney, B., and Cocks, R. (2002). Accents of guilt? Effects of regional accent, race, and crime type on attributions of guilt. *Journal of Language and Social Psychology* 21, 162–168. doi:10.1177/02627X02021002004.
- Donchin, E. (1981). Surprise!... surprise? *Psychophysiology* 18, 493–513.
- Fagyal, Z. (2003). La prosodie du français populaire des jeunes: traits héréditaires et novateurs. *Le Français aujourd'hui* 143, 47–55. Available at: http://asl.univ-montp3.fr/e41slym/Fagyal_La_prosodie_du_francais_populaire_des_jeunes_des_banlieues.pdf [Accessed January 25, 2019].

- Fagyal, Z. (2010). *Accents de banlieue. Aspects prosodiques du français populaire en contact avec les langues de l'immigration*. Paris: L'Harmattan.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., and Williams, C. J. (1995). Variability in Automatic Activation as an Unobtrusive Measure of Racial Attitudes: A Bona Fide Pipeline? *Journal of Personality and Social Psychology* 69, 1013–1027. doi:10.1037/0022-3514.69.6.1013.
- Ferragne, E., Flavier, S., and Fressard, C. (2012). ROCme! Recording of Oral Corpora Made Easy. Available at: <http://www.ddl.ish-lyon.cnrs.fr/Download/Index.asp?Langue=FR&Page=ROCme>.
- Fiebach, C. J., Schlesewsky, M., and Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research* 30, 321–338. doi:10.1023/A:1010447102554.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6, 78–84. doi:S1364661300018398 [pii].
- Grant, A., Grey, S., and van Hell, J. G. (2020). Male fashionistas and female football fans: Gender stereotypes affect neurophysiological correlates of semantic processing during speech comprehension. *Journal of Neurolinguistics* 53. doi:10.1016/j.jneuroling.2019.100876.
- Hansen, K., Steffens, M. C., Rakić, T., and Wiese, H. (2017). When appearance does not match accent: Neural correlates of ethnicity-related expectancy violations. *Social Cognitive and Affective Neuroscience* 12, 507–515. doi:10.1093/scan/nsw148.
- Helman, E., Volpert, H. I., and Simons, R. F. (2014). The N400 as an index of racial stereotype accessibility. *Social Cognitive and Affective Neuroscience* 9, 544–552. doi:10.1093/scan/nst018.
- Ibáñez, A., Haye, A., González, R., Hurtado, E., and Henríquez, R. (2009). Multi-level analysis of cultural phenomena: The role of ERPs approach to prejudice. *Journal for the Theory of Social Behaviour* 39, 81–110. doi:10.1111/j.1468-5914.2008.00391.x.
- INSEE (2003). PCS 2003 - Classification of professions and socioprofessionnal categories. Available at: <https://www.insee.fr/fr/metadonnees/pcs2003/categorieSocioprofessionnelleAgreguee/1?champRecherche=true>.
- Ito, T. A., and Bartholow, B. D. (2009). The neural correlates of race. *Trends in Cognitive Sciences* 13, 524–531. doi:10.1016/J.TICS.2009.10.002.
- Jamin, M., Trimaille, C., Gasquet-Cyrus, M., and Trimaille, C. (2006). De la convergence dans la divergence: Le cas des quartiers pluri-ethniques en France. *Journal of French Language Studies* 16, 335–356. doi:10.1017/S0959269506002559.
- Jia, L., Dickter, C. L., Luo, J., Xiao, X., Yang, Q., Lei, M., et al. (2012). Different brain mechanisms between stereotype activation and application: Evidence from an ERP study. *International Journal of Psychology* 47, 58–66. doi:10.1080/00207594.2011.580348.
- Kaan, E., Harris, A., Gibson, E., and Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15, 159–201. doi:10.1080/016909600386084.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., and Murray, R. (2007). What's new in Psychtoolbox-3. *Perception* 36, 1.

- Kurinec, C. A., and Weaver, C. A. (2019). Dialect on trial: use of African American Vernacular English influences juror appraisals. *Psychology, Crime and Law* 25, 803–828. doi:10.1080/1068316X.2019.1597086.
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology* 62, 621.
- Kutas, M., and Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology* 11, 99–116. doi:http://doi.org/10.1016/0301-0511(80)90046-0.
- Kutas, M., van Petten, C. K., and Kluender, R. (2006). “Psycholinguistics Electrified II (1994–2005),” in *Handbook of Psycholinguistics* (Elsevier), 659–724. doi:10.1016/B978-012369374-7/50018-3.
- Lattner, S., and Friederici, A. D. (2003). Talker’s voice and gender stereotype in human auditory sentence processing - Evidence from event-related brain potentials. *Neuroscience Letters* 339, 191–194. doi:10.1016/S0304-3940(03)00027-2.
- Lev-Ari, S., and Keysar, B. (2010). Why don’t we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology* 46, 1093–1096. doi:10.1016/j.jesp.2010.05.025.
- Longhi, J. (2012). “Imaginaires, représentations et stéréotypes dans la sémiotisation du mythe de la banlieue et des jeunes de banlieue,” in *Discours et sémiotisation de l’espace. Les représentations de la banlieue et de sa jeunesse* (Paris: L’Harmattan), 123–142. Available at: <https://halshs.archives-ouvertes.fr/halshs-00940249/> [Accessed January 24, 2019].
- Lopez-Calderón, J., and Luck, S. J. (2014). ERPLAB: an opensource toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience* 8, 213. doi:10.3389/fnhum.2014.00213.
- Luck, S. J. (2014). *An introduction to the event-related potential technique*. 2ème. Cambridge, MA: MIT Press.
- Mitchell, J. P., Ames, D. L., Jenkins, A. C., and Banaji, M. R. (2009). Neural correlates of stereotype application. *Journal of Cognitive Neuroscience* 21, 594–604. doi:10.1162/jocn.2009.21033.
- Molinaro, N., Su, J.-J., and Carreiras, M. (2016). Stereotypes override grammar: Social knowledge in sentence comprehension. *Brain and Language* 155–156, 36–43. doi:10.1016/J.BANDL.2016.03.002.
- Munson, B. (2007). The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and speech* 50, 125–42. doi:10.1177/00238309070500010601.
- New, B. (2006). Lexique 3 : Une nouvelle base de données lexicales. in *Actes de la Conférence Traitement Automatique des Langues Naturelles (TALN 2006)* (Louvain, Belgique).
- Osterhout, L. (1997). On the brain response to syntactic anomalies: manipulations of word position and word class reveal individual differences. *Brain and Language* 59, 494–522. doi:10.1006/brln.1997.1793.
- Osterhout, L., Bersick, M., and Mclaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition* 25, 273–285. doi:10.3758/BF03211283.

- Ouyang, G., Herzmann, G., and Zhou, C. (2011a). Residue iteration decomposition (RIDE): A new method to separate ERP components on the basis of latency variability in single trials. *Psychophysiology* 48, 1631–1647. doi:10.1111/j.1469-8986.2011.01269.x.
- Ouyang, G., Schacht, A., Zhou, C., and Sommer, W. (2013). Overcoming limitations of the ERP method with Residue Iteration Decomposition (RIDE): A demonstration in go/no-go experiments. *Psychophysiology* 50, 253–265. doi:10.1111/psyp.12004.
- Ouyang, G., Sommer, W., and Zhou, C. (2015). A toolbox for residue iteration decomposition (RIDE)—A method for the decomposition, reconstruction, and single trial analysis of event related potentials. *Journal of Neuroscience Methods* 250, 7–21. doi:10.1016/J.JNEUMETH.2014.10.009.
- Ouyang, G., Sommer, W., and Zhou, C. (2016). Reconstructing ERP amplitude effects after compensating for trial-to-trial latency jitter: A solution based on a novel application of residue iteration decomposition. *International Journal of Psychophysiology* 109, 9–20. doi:10.1016/j.ijpsycho.2016.09.015.
- Ouyang, G., Zhou, C., and Sommer, W. (2011b). RIDE toolbox. Available at: http://cns.hkbu.edu.hk/RIDE_files/Page308.htm [Accessed March 7, 2018].
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology* 81, 181–192. doi:10.1037/0022-3514.81.2.181.
- Phillips, C., Kazanina, N., and Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22, 407–428. doi:10.1016/j.cogbrainres.2004.09.012.
- Proverbio, A. M., Orlandi, A., and Bianchi, E. (2017). Electrophysiological markers of prejudice related to sexual gender. *Neuroscience* 358, 1–12. doi:10.1016/J.NEUROSCIENCE.2017.06.028.
- Ruchkin, D. S., Johnson, R., Grafman, J., Canoune, H., and Ritter, W. (1992). Distinctions and similarities among working memory processes: an event-related potential study. *Cognitive Brain Research* 1, 53–66. doi:10.1016/0926-6410(92)90005-C.
- Schad, D. J., Vasishth, S., Hohenstein, S., and Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language* 110, 104038. doi:10.1016/j.jml.2019.104038.
- Souza, R. A. de, and Oliveira, C. S. F. de (2017). Are bilingualism effects on the L1 byproducts of implicit knowledge? Evidence from two experimental tasks. *Revista de estudos da linguagem* 25, 1685. doi:10.17851/2237-2083.25.3.1685-1716.
- Steinhauer, K., and Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language* 120, 135–162. doi:10.1016/j.bandl.2011.07.001.
- Su, J. J., Molinaro, N., Gillon-Dowens, M., Tsai, P. S., Wu, D. H., and Carreiras, M. (2016). When “He” can also be “She”: An ERP study of reflexive pronoun resolution in written Mandarin Chinese. *Frontiers in Psychology* 7, 151. doi:10.3389/fpsyg.2016.00151.
- The Mathworks (2016). MATLAB.

Torre, I. (2017). The impact of voice on trust attributions. Available at: <https://pearl.plymouth.ac.uk/handle/10026.1/9858> [Accessed January 22, 2019].

van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., and Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience* 20, 580–591. doi:10.1162/jocn.2008.20054.

Wang, P., Yang, Y. P., Tan, C. H., Chen, Q. W., and Cantfort, T. (2017). Gender stereotype activation versus lexical semantic activation: An ERP study. *Journal of General Psychology* 144, 283–308. doi:10.1080/00221309.2017.1310714.

White, K. R., Crites, S. L., Taylor, J. H., and Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive and Affective Neuroscience* 4, 191–198. doi:10.1093/scan/nsp004.

APPENDIX A – Complete list of stimuli

1. Upper-class stereotypes

J'ai cours de **clavecin** à 17h ce soir.

Je fais une licence de **droit** pour être avocat.

Je suis allé boire une **coupe** avec mes amis.

Mon père est **avocat** dans un grand cabinet.

Tu as vu mon nouveau **loft** dans le 6^e ?

Je vais avoir un contrôle **fiscal** le mois prochain.

Je joue dans un **orchestre** le mercredi soir.

J'ai une réunion de la **paroisse** demain après-midi.

Ce week-end je vais faire les **boutiques**, tu veux venir ?

Où est mon nouveau **polo** vert foncé ?

J'ai mon propre **jet** pour me déplacer.

Tu préfères qu'on passe chez **Ladurée** pour les macarons ?

J'ai invité mes amis pour l'**apéritif** ce week-end.

Il faut que j'appelle mon **comptable** pour parler de mes impôts.

Mon père m'a laissé un **Picasso** en héritage.

Je l'ai vu ce matin dans le **Figaro** en première page.

J'adore jouer au **golf** avec mes associés.

Regarde le nouveau carré **Hermès** que j'ai eu à Noël.

Je vais passer le week-end à **Deauville** avec ma famille.

Ce soir je vais à un **cocktail** à l'ambassade d'Angleterre.

Je voudrais que mon frère soit **scout** dans son temps libre.

J'ai mis mon manteau en **vison** pour aller au théâtre.

J'ai envie de partir en **thalasso** en Normandie.

J'habite dans un quartier de **Versailles** avec mes parents.

Ma sœur fait de l'**équitation** le week-end.

J'ai besoin de passer au **pressing** de toute urgence.

Je vais souvent au **théâtre** avec mes amis.

Le weekend je vais dans mon **manoir** en Bretagne.

Je porte toujours des **mocassins** pour aller au travail.

Impossible, j'ai un **brunch** demain midi.

Le weekend, j'écoute de l'**opéra** à la radio.
Je dois emmener mon **lévrier** chez le vétérinaire.
J'ai acheté un **roman** à la librairie.
Je vais souvent à la **Baule** dans ma maison de vacances.
Je lave ma **rolls** toutes les semaines.
Je voyage toujours en **première** quand je prends le train.
J'adore faire du **yoga** le matin.
Je vais au lycée **international** à Paris
A Noël je mange du **caviar** de Russie.
Je bois seulement du **champagne** de grande qualité.

2. Banlieue stereotypes

Mes parents habitent à **Vitry** avec mes frères.
J'écoute toujours du **rap** dans ma voiture.
Je vais à la **mosquée** tous les vendredis.
Je vais au lycée **professionnel** dans le Val de Marne.
J'attends le versement de la **CAF** qui est en retard.
Je suis allé au Maroc en **bus** pendant l'été.
J'ai le rôle d'un grand **frère** dans la cité.
J'ai eu des cadeaux pour l'**Aïd** cette année.
Je suis allé à un concert de **Booba** la semaine dernière.
J'ai passé l'après-midi à la maison de **quartier** avec mes amis.
On va me retirer les **allocations** si je rate encore des cours.
J'habite dans un quartier **prioritaire** en Seine-Saint-Denis.
Le soir je rentre à **Grigny** vers 20h.
J'ai commencé à porter le **voile** il y a cinq ans.
Le week-end je vais **traîner** dehors avec mes amis.
Je passe le week-end à **Aulnay** chez ma tante.
J'ai eu deux heures de **colles** comme punition.
J'ai emmené ma copine au **Macdo** pour manger des frites.
Je vais cuisiner pour le **Ramadan** avec ma mère.
J'habite dans un vieux **HLM** à Aubervilliers.
Je sors toujours en **survêt** de grande marque.
J'habite dans un **ghetto** à 1h de Paris.
Je passe des heures dans le **hall** de l'immeuble.
J'achète du bœuf **halal** au supermarché.
Mon avenir c'est le **chômage** malheureusement.
Je danse le **hip-hop** en soirée.
Ma mère fait des **ménages** quelques heures par semaine.
Je fais mes courses chez **Lidl** le samedi matin.
Je suis en **décrochage** scolaire depuis la 1^e.
Mon père lit le **Coran** tous les soirs.
J'ai appelé l'assistante **sociale** pour en parler.
Je ne sors pas beaucoup de la **cité** dans laquelle j'habite.
J'ai écouté l'album de **Kaaris** qui vient de sortir.
Je viens d'un quartier de **Sarcelles** dans le Val d'Oise.
Je subis le **racisme** au quotidien.
J'habite dans cette **barre** d'immeubles qu'on voit au loin.
J'avais un jogging **Tacchini** quand j'étais ado.
Je joue souvent au **foot** le dimanche après-midi.
J'habite dans des **tours** en banlieue parisienne.
J'étais dans un collège de **ZEP** à mauvaise réputation.

3. Fillers

Les frais sont **élevés** / ***furieux** dans cette université.
Ce livre est **ancien** / ***ami** et précieux.
Ce plat est vraiment **chaud** / ***fier**, fais attention.

Le placard est **étroit** / *adulte et peu commode.
J'ai vu ce film en **noir** / *flou et blanc.
Le signal est **faible** / *sage dans la cuisine.
C'est un homme **sincère** / *mental et toujours franc.
Il portait un costume **élégant** / *amical à la conférence.
Les tigres sont **rares** / *droits à l'état sauvage.
Cet investissement sera **utile** / *charmant pour la communauté.
C'est une salade **saine** / *lâche et très originale.
Le trajet est un peu **long** / *pur pour le faire ce week-end.
Ici, fumer est **interdit** / *habillé sous peine d'amende.
Ce soda est bien **frais** / *gris grâce aux glaçons.
Son frère est **marié** / *urgent depuis deux ans.
Ce collier est **unique** / *jaloux et hors de prix.
Ce film est **nul** / *vaste du début à la fin.
La situation est **compliquée** / *généreuse au Royaume Uni.
Ils ont un château **splendide** / *timide en Bretagne.
La réunion est **publique** / *atteinte donc tu peux venir.
Cet enfant est **curieux** / *désert et intelligent.
Le goût est **étonnant** / *illégal mais pas désagréable.
L'eau est très **claire** / *grosse dans cette fontaine.
Ces vêtements sont trop **vieux** / *tards pour être mettables.
Je voudrais une chaise **rouge** / *triste pour mon salon.
Mon collègue n'est pas **positif** / *douloureux du tout, c'est insupportable.
J'ai mangé du pain **délicieux** / *courageux au restaurant.
Les manifestants sont **volontaires** / *fédéraux et en colère.
Ce tableau est trop **laid** / *ivre pour que je l'achète.
Cette punition est **sévère** / *mourante pour un enfant.
Ce mur est trop **haut** / *mort pour voir par-dessus.
La qualité est **supérieure** / *adorable dans ce magasin.
Mon chat est **doux** / *nu et affectueux.
Son appartement est **sale** / *creux en permanence.
Le café est **gratuit** / *sauvage avec le menu.
Ce bureau a l'air **solide** / *acide et confortable.
Il avait une vision **moderne** / *assise de la société.
Son emplacement est **idéal** / *imbécile pour se baigner.
Ce pull est **affreux** / *inquiet, il faut le jeter.
Ce vélo est vraiment **lourd** / *tendre pour un vélo de course.
Ce n'est pas aussi **simple** / *capable qu'il n'y paraît.
Les tiroirs étaient **ouverts** / *joyeux quand je suis rentré.
Ce sac est très **léger** / *prudent donc je l'emporte partout.
Ses notes sont **meilleures** / *dernières en maths qu'en français.
J'ai mis mon t-shirt **vert** / *cru pour l'occasion.
Cette pièce est trop **sombre** / *brave pour pouvoir y lire.
J'ai entendu un bruit **sourd** / *saint dans le grenier.
Ce dessert est trop **amer** / *serré pour que je l'apprécie.
Je rends visite à un **ami** / *hasard de mes parents.
Elle a parlé à la **classe** / *règle du prochain contrôle.
La décision du **ministre** / *manteau est très critiquée.

Je l'ai vu après son **opération** / ***information** du genou.
J'ai expliqué au **patient** / ***regard** quel était son problème.
Ce chien est d'une **race** / ***feuille** très recherchée.
Je mange des **pâtes** / ***cendres** presque tous les soirs.
Je fais partie d'une **équipe** / ***fièvre** de basket.
Les carottes de mon **lapin** / ***salon** commencent à moisir.
J'ai cassé une **bouteille** / ***pensée** de champagne.
J'ai mal aux **oreilles** / ***étoiles** à cause du bruit.
Il s'est adressé à la **foule** / ***goutte** d'un ton ferme.
Il a fait son **devoir** / ***espoir** de maths pour demain.
J'ai mal à la **main** / ***porte** depuis une semaine.
Je suis allé à une **réunion** / ***caméra** la semaine dernière.
Tu viens au **cinéma** / ***souvenir** ce week-end ?
Il travaille dans des **champs** / ***goûts** de betteraves à sucre.
Il a trouvé une **solution** / ***maladie** à notre problème.
Elle a lavé la **cuisine** / ***semaine** de fond en comble.
J'ai mangé une **glace** / ***pierre** pour le goûter.
Elle travaille dans une **entreprise** / ***habitude** de construction automobile.
Le héros est un **sorcier** / ***contrat** de onze ans.
On est mieux sous les **arbres** / ***jeux** quand il y a du soleil.
Je vais prendre un **thé** / ***tort** à emporter.
Elle a donné son **avis** / ***mari** sur la question.
J'ai mal aux **yeux** / ***jours** à cause du soleil.
Il a étendu son **empire** / ***auteur** sur toute l'Europe.
Il raconte des **histoires** / ***monnaies** à ses enfants.
Il a discuté avec sa **tante** / ***ferme** des cadeaux de Noël.
Ça se passe dans un **terrain** / ***destin** vague en ville.
J'aime rentrer dans mon **village** / ***rapport** pendant les vacances.
Elle a observé l'**oiseau** / ***ennui** par la fenêtre.
La performance de ce **moteur** / ***besoin** est impressionnante.
Il a servi le **vin** / ***doigt** avant le repas.
J'ai échangé mon **sac** / ***don** contre un plus grand.
J'ai enfilé une **chaussure** / ***adresse** pour l'essayer.
L'usine va acheter une **machine** / ***colère** pour faire cette tâche.
Il portait un **chapeau** / ***genou** haut de forme.
Elle a acheté une **île** / ***reine** déserte dans le pacifique.
Il faudrait baisser le **prix** / **chef** de cet article.
Tu peux me rendre un **service** / ***service** s'il te plait ?
Elle a nourri ses **chiens** / ***roses** avant de partir.
Il prend facilement **contact** / ***sommeil** avec les gens.
Ils ont embauché une **bonne** / ***drogue** pour s'occuper des enfants.
On entend bien le **vent** / ***haut** dans cette cabane.
Elle rentre à la **maison** / ***personne** après 20h.
Je préfère garder mes **distances** / ***régions** avec ces gens-là.
Il a pris ses **bagages** / **humain** en cabine.

APPENDIX B – Full model reports

1. Behavioral results

1.1 Congruence effect on image response time and accuracy for semantic mismatches

	β	<i>SE</i>	<i>df</i>	<i>t/z</i>	<i>Significance</i>
RESPONSE TIME					
<i>Log10(RT) ~ Congruence + (1 Participant) + (1 item)</i>					
Intercept	2.54	0.02	26	136.27	
Congruence	0.03	0.009	85	3.67	$\chi^2(1) = 12.82, p = .0003$
ACCURACY					
<i>Accuracy ~ Congruence + (Congruence Participant) + (Congruence item), family=binomial(link=logit)</i>					
Intercept	3.41	0.25		13.52	
Congruence	-0.71	0.44	8	-1.62	$\chi^2(1) = 2.64, p = .10$

1.2 Congruence effect on image response time and accuracy for stereotypical mismatches

	β	<i>SE</i>	<i>df</i>	<i>t/z</i>	<i>Significance</i>
RESPONSE TIME					
<i>Log10(RT) ~ Congruence + Accent + Congruence:Accent + (1 Participant) + (Accent item)</i>					
Intercept	2.51	0.02	26	130.13	
Congruence	-0.001	0.007	70	-0.20	$\chi^2(1) = 0.12, p = .73$
Accent	-0.007	0.007	70	-0.90	$\chi^2(1) = 1.01, p = .32$
Congruence:Accent	0.02	0.02	76	1.13	$\chi^2(1) = 1.26, p = .26$
ACCURACY					
<i>Accuracy ~ Congruence + Accent + Congruence:Accent + (Congruence Participant) + (Accent item), family=binomial(link=logit)</i>					
Intercept	2.64	0.16		16.83	
Congruence	-0.03	0.21	7	-0.13	$\chi^2(1) = 1.60, p = .21$
Accent	-0.18	0.20	7	-0.86	$\chi^2(1) = 0.47, p = .49$
Congruence:Accent	-0.36	0.43	8	-0.85	$\chi^2(1) = 0.44, p = .51$

1.3 Response time and accuracy for crime-related images

	β	<i>SE</i>	<i>df</i>	<i>t/z</i>	<i>Significance</i>
RESPONSE TIME					
<i>Log10(RT) ~ Accent + (1 Participant) + (1 item)</i>					
Intercept	5.84	0.05	29	119.08	
Accent	0.02	0.02	244	0.88	$\chi^2(1) = 0.77, p = .38$
ACCURACY					
<i>Accuracy ~ Accent + (1 Participant) + (1 item), family=binomial(link=logit)</i>					
Intercept	4.21	0.49		8.67	
Accent	0.59	0.42	4	1.38	$\chi^2(1) = 1.96, p = .16$

2. EEG results – Semantic mismatches

NOTE: In the following tables, the significance of each of the main effect and of the interaction is reported from log-likelihood model comparisons. Estimates are reported from a summary of the model. If the interaction was significant, a nested contrasts model (Mean amplitude ~ Region/Congruence + random-effects structure) was run to investigate in which regions the Congruence effect was significant. The reported estimates and p values for the levels of the Congruence:Region interaction come from that model.

2.1 A priori defined time-window for N400 effect: 300-700 ms

LATERAL ANALYSES

Mean amplitude 300-700 ms ~ Congruence + Region + Congruence:Region + (Congruence|Participant)

	β	SE	df	t	p
Intercept	0.17	0.18	24	0.96	
Congruence	$\chi^2(1) = 3.69, p = .07$				
Congruence	-0.59	0.31	24	-1.90	
Region	$\chi^2(5) = 279.67, p < .0001$ ***				
Anterior Right	-0.06	0.08	1968	-0.80	
Central Left	-0.15	0.07	1968	-2.09	
Central Right	-0.28	0.07	1968	-3.92	
Posterior Left	-0.82	0.08	1968	-10.58	
Posterior Right	-1.00	0.08	1968	-12.83	
Congruence:Region	$\chi^2(5) = 16.75, p = .005$ **				
Congruence: Anterior Left	-0.54	0.33	29	-1.64	.11
Congruence: Anterior Right	-0.38	0.33	29	-1.16	.25
Congruence: Central Left	-0.88	0.32	27	-2.73	.01*
Congruence: Central Right	-0.72	0.32	27	-2.24	.03*
Congruence: Posterior Left	-0.57	0.33	29	-1.74	.09.
Congruence: Posterior Right	-0.47	0.33	29	-1.44	.16

MIDLINE ANALYSES

Mean amplitude 300-700 ms ~ Congruence + Region + Congruence:Region + (Congruence+Region|Participant)

	β	SE	df	t	p
Intercept	-0.04	0.24	24	-0.15	
Congruence	$\chi^2(1) = 1.28, p = .26$				
Congruence	-0.82	0.42	25	-1.94	
Region	$\chi^2(2) = 35.84, p < .0001$ ***				
Region Central	-0.42	0.25	24	-1.69	
Region Posterior	-1.84	0.25	24	-7.30	
Congruence:Region	$\chi^2(2) = 7.10, p = .029$ *				
Congruence: Anterior	-1.13	0.43	27	-2.61	.015 *
Congruence: Central	-0.98	0.43	26	-2.30	.030 *
Congruence: Posterior	-0.65	0.43	27	-1.51	0.14

2.2 A priori defined time-window for late effects: 700-900 ms

LATERAL ANALYSES					
<i>Mean amplitude 700-900 ms ~ Congruence + Region + Congruence:Region + (Congruence Participant)</i>					
	β	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	0.14	0.15	24	0.98	
Congruence	$\chi^2(1) = 0.22, p < .64$				
Congruence	-0.12	0.30	24	-0.41	
Region	$\chi^2(5) = 145.84, p < .0001$ ***				
Anterior Right	-0.18	0.07	1968	-2.68	
Central Left	0.14	0.07	1968	2.18	
Central Right	-0.01	0.06	1968	-0.23	
Posterior Left	-0.43	0.07	1968	-6.21	
Posterior Right	-0.52	0.07	1968	-7.59	
Congruence:Region	$\chi^2(5) = 159.87, p < .0001$ ***				
Congruence: Anterior Left	-0.76	0.31	29	-2.46	.020 *
Congruence: Anterior Right	-0.45	0.31	29	-1.44	.16
Congruence: Central Left	-0.46	0.31	27	-1.52	.14
Congruence: Central Right	-0.04	0.31	27	-0.13	.90
Congruence: Posterior Left	0.27	0.31	29	0.87	.39
Congruence: Posterior Right	0.72	0.31	29	2.32	.028 *

MIDLINE ANALYSES					
<i>Mean amplitude 700-900 ms ~ Congruence + Region + Congruence:Region + (Congruence Participant)</i>					
	β	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	0.07	0.18	25	0.40	
Congruence	$\chi^2(1) = 0.19, p = .67$				
Congruence	0.19	0.35	26	0.53	
Region	$\chi^2(2) = 42.12, p < .0001$ ***				
Region Central	0.07	0.16	288	0.44	
Region Posterior	-1.10	0.18	288	-5.71	
Congruence:Region	$\chi^2(2) = 31.99, p < .0001$ ***				
Congruence: Anterior	-0.78	0.37	33	-2.11	.043 *
Congruence: Central	-0.21	0.36	29	-0.60	.55
Congruence: Posterior	0.58	0.37	33	1.58	.12

2.3 Supplementary analyses in the 0-300 ms window

LATERAL ANALYSES					
<i>Mean amplitude 0-300 ms ~ Congruence + Region + Congruence:Region + (Congruence Participant)</i>					
	β	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	0.28	0.17	24	1.63	
Congruence	$\chi^2(1) = 1.13, p = .29$				
Congruence	-0.24	0.25	24	-0.98	
Region	$\chi^2(5) = 11.18, p = .048$ *				
Anterior Right	0.03	0.08	1968	0.37	
Central Left	0.17	0.08	1968	2.43	
Central Right	0.17	0.07	1968	2.38	
Posterior Left	0.03	0.08	1968	0.39	
Posterior Right	0.09	0.08	1968	1.17	

Congruence:Region	$\chi^2 (5) = 58.22, p < .0001$ ***				
Congruence: Anterior Left	-0.47	0.27	33	-1.77	.086 .
Congruence: Anterior Right	-0.61	0.27	33	-2.27	.030 *
Congruence: Central Left	-0.40	0.26	29	-1.52	.14
Congruence: Central Right	-0.42	0.26	29	-1.64	.11
Congruence: Posterior Left	0.18	0.27	33	0.68	.50
Congruence: Posterior Right	0.27	0.27	33	1.01	.32

MIDLIN ANALYSES

Mean amplitude 0-300 ms ~ Congruence + Region + Congruence:Region + (Congruence|Participant)

	β	SE	df	t	p
Intercept	0.54	0.22	25	2.44	
Congruence	$\chi^2 (1) = 1.95, p = .16$				
Congruence	-0.13	0.32	26	-0.41	
Region	$\chi^2 (2) = 12.19, p = .002$ **				
Region Central	0.49	0.17	288	2.89	
Region Posterior	-0.06	0.19	288	-0.35	
Congruence:Region	$\chi^2 (2) = 27.11, p < .0001$ ***				
Congruence: Anterior	-1.05	0.35	36	-3.01	.005 **
Congruence: Central	-0.52	0.33	30	-1.57	.13
Congruence: Posterior	0.26	0.35	36	0.75	.46

3. EEG results – Stereotypical mismatches

NOTE: In the following tables, the significance of each of the main effect and of the interactions is reported from log-likelihood model comparisons. Estimates are reported from a summary of the model. If the two-way interaction between Congruence and Region was significant in one of the submodels or in the original full model, a nested contrasts model (Mean amplitude ~ Region/Congruence + random-effects structure) was run to investigate in which regions the Congruence effect was significant. The reported estimates and *p* values for the levels of the Congruence * Region interaction come from that model.

3.1 A priori defined time-window for N400: 300-700 ms

LATERAL ANALYSES					
<i>Mean amplitude 300-700 ms ~ Congruence * Region + (Congruence Participant)</i>					
	β	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	-0.17	0.08	24	-2.09	
Congruence	$\chi^2(1) = 0.88, p = .35$				
Congruence	0.06	0.09	25	0.67	
Region	$\chi^2(5) = 140.28, p < .0001$ ***				
Anterior Right	-0.23	0.10	1968	-2.34	
Central Left	-0.21	0.09	1968	-2.40	
Central Right	-0.44	0.09	1968	-5.03	
Posterior Left	-0.68	0.10	1968	-7.04	
Posterior Right	-1.02	0.10	1968	-10.50	
Congruence:Region	$\chi^2(5) = 64.74, p < .0001$ ***				
Congruence: Anterior Left	-0.70	0.16	199	-4.50	<.0001***
Congruence: Anterior Right	-0.10	0.16	199	-0.63	0.53
Congruence: Central Left	-0.48	0.13	113	-3.59	.0005 ***
Congruence: Central Right	0.009	0.13	113	0.06	0.95
Congruence: Posterior Left	0.35	0.16	199	2.26	.025 *
Congruence: Posterior Right	0.56	0.16	199	3.57	.0005 ***

MIDLINE ANALYSES					
<i>Mean amplitude 300-700 ms ~ Congruence * Region + (Congruence Participant)</i>					
	β	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	-0.40	0.15	27	-2.71	
Congruence	$\chi^2(1) = 0.41, p = .52$				
Congruence	0.24	0.17	33.68	1.39	
Region	$\chi^2(2) = 83.59, p < .0001$ ***				
Region Central	-1.02	0.19	288	-5.41	
Region Posterior	-2.13	0.21	288	-10.24	
Congruence:Region	$\chi^2(2) = 24.49, p < .0001$ ***				
Congruence: Anterior	-0.81	0.23	90	-3.56	<.001***
Congruence: Central	-0.07	0.20	56	-0.35	0.73
Congruence: Posterior	0.55	0.23	90	2.44	0.02 *

3.2 A priori defined time-window for late effects: 700-900 ms

LATERAL ANALYSES					
<i>Mean amplitude 700-900 ms ~ Congruence * Region * SES + (Congruence Participant)</i>					
	β	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	-0.004	0.006	24	-0.08	
Congruence	$\chi^2(1) = 0.18, p = .67$				

Congruence	0.06	0.12	24	0.49	
Region	$\chi^2(5) = 117.19, p < .0001$ ***				
Anterior Right	-0.32	0.08	1968	-4.29	
Central Left	0.20	0.07	1968	2.94	
Central Right	-0.08	0.07	1968	-1.14	
Posterior Left	-0.18	0.08	1968	-2.42	
Posterior Right	-0.49	0.08	1968	-6.43	
Congruence:Region	$\chi^2(5) = 42.80, p < .0001$ ***				
Congruence: Anterior Left	-0.32	0.15	70	-2.10	.040 *
Congruence: Anterior Right	-0.15	0.15	70	-0.98	.33
Congruence: Central Left	-0.09	0.14	49	-0.67	.51
Congruence: Central Right	0.09	0.14	49	0.67	.51
Congruence: Posterior Left	0.30	0.15	70	1.95	.055 .
Congruence: Posterior Right	0.52	0.15	70	3.38	.001 **

MIDLINE ANALYSES

*Mean amplitude 700-900 ms ~ Congruence * Region * Background + (Congruence|Participant)*

	β	SE	df	t	p
Intercept	0.11	0.11	28	1.06	
Congruence	$\chi^2(1) = 0.63, p = .43$				
Congruence	0.37	0.18	29	2.02	
Region	$\chi^2(2) = 26.48, p < .0001$ ***				
Region Central	-0.10	0.15	288	-0.68	
Region Posterior	-0.80	0.16	288	-4.87	
Congruence:Region	$\chi^2(2) = 16.35, p = .0003$ ***				
Congruence: Anterior	-0.37	0.22	58	-1.70	.095 .
Congruence: Central	0.23	0.20	41	1.17	.25
Congruence: Posterior	0.50	0.22	58	2.33	.023 *

3.3 Additional analysis: 100-300 ms

LATERAL ANALYSES

*Mean amplitude 100-300 ms ~ Congruence * Region * SES + (Congruence|Participant)*

	β	SE	df	t	p
Intercept	-0.12	0.08	24	-1.61	
Congruence	$\chi^2(1) = 0.05, p = .82$				
Congruence	0.04	0.11	24	0.34	
Region	$\chi^2(5) = 9.42, p = .093$				
Anterior Right	-0.13	0.09	1968	-1.40	
Central Left	-0.19	0.08	1968	-2.28	
Central Right	-0.14	0.08	1968	-1.75	
Posterior Left	-0.24	0.09	1968	-2.67	
Posterior Right	-0.23	0.09	1968	-2.55	
Congruence:Region	$\chi^2(5) = 55.95, p < .0001$ ***				
Congruence: Anterior Left	-0.48	0.16	106	-2.96	.004 **
Congruence: Anterior Right	-0.21	0.16	106	-1.30	.20
Congruence: Central Left	-0.22	0.15	67	-1.51	.13
Congruence: Central Right	0.12	0.15	67	0.86	.40
Congruence: Posterior Left	0.38	0.16	106	2.34	.021 *
Congruence: Posterior Right	0.64	0.16	106	3.92	.0002 ***

MIDLINE ANALYSES

Mean amplitude 100-300 ms ~ Congruence * Region * SES + (Congruence|Participant)

	β	SE	df	t	p
Intercept	-0.15	0.16	26	-0.99	
Congruence	$\chi^2(1) = 0.85, p = .36$				
Congruence	0.50	0.19	32	2.72	
Region	$\chi^2(2) = 10.29, p = .006^{**}$				
Region Central	-0.24	0.18	288	-1.36	
Region Posterior	-0.65	0.20	288	-3.34	
Congruence:Region	$\chi^2(2) = 25.67, p < .0001^{***}$				
Congruence: Anterior	-0.55	0.23	75	-2.37	.021 *
Congruence: Central	0.21	0.21	49	1.01	.32
Congruence: Posterior	0.80	0.23	75	3.44	.0009 ***

4. EEG results – Stereotypical mismatches with SES background

NOTE: In the following tables, the significance of each of the main effect and of the interactions is reported from log-likelihood model comparisons. Estimates are reported from a summary of the model. If the three-way interaction was significant, separate follow-up models were run for each level of the Socio-economic status (SES) variable (Upper class / not upper class). If the two-way interaction between Congruence and Region was significant in one of the submodels or in the original full model, a nested contrasts model (Mean amplitude ~ Region/Congruence + random-effects structure) was run to investigate in which regions the Congruence effect was significant. The reported estimates and *p* values for the levels of the Congruence:Region interaction come from that model.

3.4 A priori defined time-window for N400: 300-700 ms

LATERAL ANALYSES

Mean amplitude 300-700 ms ~ Congruence * Region * SES + (Congruence|Participant)

	β	SE	df	t	p
Intercept	-0.18	0.08	24	-2.13	
Congruence	$\chi^2(1) = 0.88, p = .35$				
Congruence	-0.05	0.09	25	-0.50	
Region	$\chi^2(5) = 140.28, p < .0001^{***}$				
Anterior Right	-0.22	0.10	1968	-2.24	
Central Left	-0.25	0.09	1968	-2.91	
Central Right	-0.50	0.09	1968	-5.51	
Posterior Left	-0.73	0.10	1968	-7.36	
Posterior Right	-1.03	0.10	1968	-10.39	
SES	$\chi^2(1) = 0.31, p = .58$				
SES	-0.06	0.17	24	-0.38	
Congruence:Region	$\chi^2(5) = 64.74, p < .0001^{***}$				
Congruence: Anterior Left	-0.70	0.16	199	-4.50	<.0001***
Congruence: Anterior Right	-0.10	0.16	199	-0.63	0.53
Congruence: Central Left	-0.48	0.13	113	-3.59	.0005 ***
Congruence: Central Right	0.009	0.13	113	0.06	0.95
Congruence: Posterior Left	0.35	0.16	199	2.26	.025 *
Congruence: Posterior Right	0.56	0.16	199	3.57	.0005 ***
Congruence:SES	$\chi^2(1) = 0.32, p = .57$				
Region:SES	$\chi^2(5) = 14.13, p = .015 *$				
Congruence:Region:SES	$\chi^2(5) = 6.75, p = .24$				

MIDLINE ANALYSES

Mean amplitude 300-700 ms ~ Congruence * Region * SES + (Congruence|Participant)

	β	SE	df	t	p
Intercept	-0.46	0.15	27	-3.06	
Congruence	$\chi^2(1) = 0.41, p = .52$				
Congruence	0.22	0.18	33.69	1.23	
Region	$\chi^2(2) = 83.59, p < .0001$ ***				
Region Central	-1.02	0.19	288	-5.41	
Region Posterior	-2.13	0.21	288	-10.24	
SES	$\chi^2(1) = 1.33, p = .25$				
SES	-0.44	0.30	27	-1.47	
Congruence:Region	$\chi^2(2) = 24.49, p < .0001$ ***				
Congruence: Anterior	-0.81	0.23	90	-3.56	<.001***
Congruence: Central	-0.07	0.20	56	-0.35	0.73
Congruence: Posterior	0.55	0.23	90	2.44	0.02 *
Congruence:SES	$\chi^2(1) = 0.35, p = .55$				
Region:SES	$\chi^2(2) = 4.01, p = .13$				
Congruence:Region:SES	$\chi^2(2) = 0.02, p = .99$				

3.5 A priori defined time-window for late effects: 700-900 ms

LATERAL ANALYSES

Mean amplitude 700-900 ms ~ Congruence * Region * SES + (Congruence|Participant)

	β	SE	df	t	p
Intercept	0.003	0.006	24	0.06	
Congruence	$\chi^2(1) = 0.18, p = .67$				
Congruence	0.11	0.11	24	0.95	
Region	$\chi^2(5) = 117.19, p < .0001$ ***				
Anterior Right	-0.34	0.08	1968	-4.31	
Central Left	0.16	0.07	1968	2.24	
Central Right	-0.12	0.07	1968	-1.65	
Posterior Left	-0.20	0.08	1968	-2.56	
Posterior Right	-0.50	0.08	1968	-6.45	
SES	$\chi^2(1) = 0.03, p = .85$				
SES	0.06	0.11	24	0.54	
Congruence:Region	$\chi^2(5) = 42.80, p < .0001$ ***				
Congruence: Anterior Left	-0.32	0.15	70	-2.10	.040 *
Congruence: Anterior Right	-0.15	0.15	70	-0.98	.33
Congruence: Central Left	-0.09	0.14	49	-0.67	.51
Congruence: Central Right	0.09	0.14	49	0.67	.51
Congruence: Posterior Left	0.30	0.15	70	1.95	.055 .
Congruence: Posterior Right	0.52	0.15	70	3.38	.001 **
Congruence:SES	$\chi^2(1) = 3.39, p = .066 .$				
Region:SES	$\chi^2(5) = 9.39, p = .099 .$				
Congruence:Region:SES	$\chi^2(5) = 9.27, p = .099 .$				

MIDLINE ANALYSES

Mean amplitude 700-900 ms ~ Congruence * Region * Background + (Congruence|Participant)

	β	SE	df	t	p
Intercept	0.09	0.11	28	0.85	
Congruence	$\chi^2(1) = 0.63, p = .43$				
Congruence	0.41	0.19	29	2.22	
Region	$\chi^2(2) = 26.48, p < .0001$ ***				

Region Central	-0.19	0.15	288	-1.22	
Region Posterior	-0.83	0.17	288	-4.93	
SES	$\chi^2(1) = 0.24, p = .63$				
SES	-0.16	0.22	28	-0.72	
Congruence:Region	$\chi^2(2) = 16.35, p = .0003$ ***				
Congruence: Anterior	-0.37	0.22	58	-1.70	.095 .
Congruence: Central	0.23	0.20	41	1.17	.25
Congruence: Posterior	0.50	0.22	58	2.33	.023 *
Congruence:SES	$\chi^2(1) = 0.61, p = .44$				
Region:SES	$\chi^2(2) = 5.27, p = .07$				
Congruence:Region:SES	$\chi^2(2) = 1.23, p = .54$				