IMPLICIT VS. EXPLICIT PERCEPTION OF FRENCH OPTIONAL LIAISON AS A MARKER OF FORMALITY

Megan M. Dailey, Sharon Peperkamp

Laboratoire de Sciences Cognitives et Psycholinguistique (ENS-PSL, CNRS, EHESS)
megan.dailey@ens.psl.eu, sharon.peperkamp@ens.psl.eu

ABSTRACT

This study examines the effect of explicit awareness (or, *noticing*) of socially-indexed variation on the generation of social inferences during speech perception. In a matched-guise task, French listeners associated sentences that contained optional liaison—a sociolinguistic variable at the level of *stereotype* in French—more frequently with 'formal language' than with 'everyday language'. This effect was larger for participants who reported noticing the optional liaisons than for participants who did not.

These results demonstrate that while noticing variation in context is not a necessary condition for social inferences to be generated, language users' subjective experience of socially-conditioned variation influences the frequency with which social inferences arise, which may ultimately lead to individual differences in the representation of the same socially-indexed variation.

Keywords: sociophonetics, speech perception, awareness, individual differences, French liaison.

1. INTRODUCTION

Social inferences about a speaker or context arise rapidly, automatically, and implicitly during speech perception [1]. At the same time, language users have explicit knowledge or awareness of some, but not all socially-indexed variation in the speech signal [2], and the extent of this knowledge varies by sociolinguistic variable (i.e., an association of linguistic form and social category). Labov [2] defines a scale of awareness of sociolinguistic variables that distinguishes *indicators*, of which language users have no awareness, *markers*, of which they have implicit awareness, in that they generate social inferences upon encountering a marker without explicit awareness, and *stereotypes*, of which they have explicit awareness and which are often the subject of metalinguistic commentary.

There is a large body of work dedicated to describing language users' knowledge of sociolinguistic variation [3]; additionally, evidence from sociophonetic perception studies has demonstrated that social information given to listeners about a speaker or context can bias speech perception [4-5]. Yet little is known about the effect of explicit awareness of socially-indexed variation on social inferences made during language processing (described as *noticing* by Schmidt [6]).

In a study on noticing using a self-paced reading task, Squires [7] compared listener awareness of standard, dialectal and unattested subject-verb agreement forms. Results showed that the different-dialect form [singular-NP + *don't*], a variable at the level of *stereotype* in American English, impeded reaction times of participants who reported noticing the form in a follow-up questionnaire to a greater extent than those who did not report noticing the form, though both groups exhibited slower reaction times to non-standard forms than to standard forms. Therefore, while dialectal and unattested variation affected the performance of both groups, it was noticing that determined the speed at which variants were processed.

In this paper, we aim to quantify the effect of noticing on the generation of social inferences related to it. Our case study involves French liaison, a complex external sandhi phenomenon in which certain word-final consonants are produced only when the following word starts with a vowel. In French, liaison is obligatory in some contexts (e.g., between an article and a following noun, as in *les‿idées* [le**z**ide] 'the ideas'), and optional in others (e.g., between a plural noun and a following adjective, as in *idées‿intéressantes* [ide**z**ɛ̃teʁesɑ̃t] or [ideɛ̃teʁesɑ̃t] 'interesting ideas'). The variation induced by optional liaison is socially structured: corpus studies of French have found higher rates of optional liaison in formal speech (radio news broadcast) than in casual speech (conversations between friends) [8-9]. These empirical observations back a theoretical hierarchy of discourse contexts in which different optional liaisons would most likely be deployed [10].

Optional liaison is a sociolinguistic variable at the level of *stereotype* in Metropolitan French. It is associated with a hyper-standard, prestige variety of French learned in school and linked to political, legal, academic and business language [11]. Although optional liaison is a robustly studied phenomenon, there is to our knowledge no experimental data showing that French listeners systematically link the

presence of optional liaison with formality[1]. The goal of this study, then, is two-fold: we examine, first, whether French listeners reliably associate optional liaison to formality during speech perception, and, second, whether their noticing of optional liaison mediates the strength of this association. To do so, we employ a matched-guise paradigm [13] and a post-experiment questionnaire probing explicit awareness.

2. METHODS

2.1. Materials

Stimuli were 17 short, descriptive French sentences, each with a single optional liaison environment. They had simple syntactic structure and contained neutral observations or well-known facts. Sentences were constructed such that obvious lexical and other phonetic cues that signal discourse (in)formality in French were absent. In particular, they contained no environments for schwa-deletion or post-obstruent liquid deletion (both of which signal casual speech in French [14]), no singular second-person pronouns (*tu* or *vous*, depending on relational proximity), no first-person plural pronouns (*on* or *nous*, depending on level of discourse formality), and – with one exception[2] – no negations (e.g., *pas* or *ne…pas* 'not', depending on level of discourse formality). An example, with the optional liaison environment underlined and bolded, is shown in (1). A full list of sentences can be found in the Appendix.

(1) *Les plat**s**‿**i**taliens sont très populaires.*
Italian dishes are very popular.

Sentences were recorded in a soundproof booth by a male native speaker of French with state-of-the-art recording equipment. The speaker was instructed to produce each sentence both with and without liaison, at a normal speaking rate and by keeping the productions as similar as possible for all aspects except liaison. Both guises were recorded 2-3 times, and the best takes were selected for subsequent manipulation in Praat [15].

Sentences were first scaled to match in average intensity, and then manipulated in the following manner: For each sentence, the two-word sequence that constituted the optional liaison environment was spliced out of both guises. Pitch contours and duration were then matched across guises, and the resulting sequences were both spliced back into the sentence originally produced in the guise without liaison. As a result, the two guises of a sentence differed only in the presence or absence of liaison in the crucial two-word sequence.

Two counterbalanced stimulus lists were created, varying which sentences were presented with and without liaison in each, such that no sentence appeared in the same list twice. The number of sentences with and without liaison was also balanced across lists, such that in the first list, there were 8 sentences with liaison and 9 sentences without, and vice versa for the second list.

2.2. Participants

Participants were 109 native speakers of Metropolitan French (age range: 19-60, median: 26), located in France and recruited on Prolific (www.prolific.co). Data from three additional participants who were speakers of other varieties of French were collected but not analyzed.

2.3. Procedure

The experiment was run on the experiment-hosting platform Labvanced [16]. Participants were instructed to sit in a quiet room and use headphones to complete the task. They were randomly assigned to one of the two lists. Within the relevant list, sentences were presented in a random order.

In each trial, participants listened to a sentence and subsequently provided a binary judgment on its formality: *langage soutenu* – "formal language" or *langage courant* – "everyday language". In order to minimize the effect of semantic content on formality judgments, participants were instructed to pay more attention to what each sentence sounded like than to its meaning.

At the end of the experiment, participants completed a questionnaire on their experience with the task as well as their language background and demographic characteristics. Two questions, shown in (2), probed participants' explicit awareness of optional liaison in the sentence stimuli.

(2) *a. Quelle stratégie avez-vous utilisée pour faire vos choix ?*
What strategy did you use to make your choices?
b. Y a-t-il un aspect des phrases qui a particulièrement attiré votre attention ?
Was there an aspect of the sentences that attracted your attention in particular?

3. RESULTS

Participants were first assigned to one of two groups based on their mention of optional liaison in responses to the two relevant questions in the post-experiment questionnaire. Those that included the word *liaison* in at least one of their responses to the

two relevant questions above or gave an example of an optional liaison they heard in the experiment were coded as Mentioners (n = 32), and those that did not were coded as Non-Mentioners (n = 77).

Formality ratings were coded as 1 for *langage soutenu* 'formal language' and as 0 for *langage courant* 'everyday language'. The overall mean formality rating was 0.31; hence, sentences were rated as 'everyday language' more than as 'formal language'. Table 1 reports mean formality ratings by Liaison and Mention status, and Figure 1 shows individual difference scores of formality ratings for sentences with minus sentences without optional liaison.

Table 1. Mean formality ratings by Liaison and Mention status. Standard errors are shown in parentheses.

Mention	Liaison	No Liaison
Yes	0.55 (0.09)	0.24 (0.05)
No	0.32 (0.08)	0.23 (0.05)

We performed a logistic mixed effects regression with fixed effects Liaison, Mention, and their interaction, random intercepts for Participant and Sentence pair, and a random by-sentence slope of Mention. The analysis was carried out with the lme4 package [17] in the R environment [18]; significance was established by means of model comparison and a likelihood ratio test, and the random structure was the maximal one that allowed for convergence while avoiding overfitting. Results revealed main effects of Liaison (β = 0.60, SE = 0.07, z = 8.83, $Chi^2(1)$ = 83.5, $p < 0.001$) and Mention (β = 0.37, SE = 0.17, z = 2.17, $Chi^2(1)$ = 4.58, $p < 0.04$), and a significant interaction between the two (β = 0.29, SE = 0.07, z = 4.19, $Chi^2(1)$ = 17.8, $p < 0.001$).

Sentences with liaison were rated as more formal than sentences without liaison, and Mentioners rated sentences as more formal overall than Non-Mentioners. Further, the difference in rating for sentences with and without liaison was bigger for Mentioners than for Non-Mentioners.

4. DISCUSSION

In a matched-guise task, French adult listeners were sensitive to the presence of optional liaison as a sociolinguistic cue, linking it to formal language. In other words, the presence of optional liaison biased interpretations of sentence formality even in the absence of noticing, indicating that French listeners have implicit knowledge of optional liaison as a marker of formality. Furthermore, listeners who reported noticing optional liaison in the sentences they listened to made this association more frequently than those who did not. Based on our finding that noticing socially-conditioned variation leads to more frequent generation of social inferences related to it, we hypothesize that listeners who tend to notice socially-indexed variation during speech perception

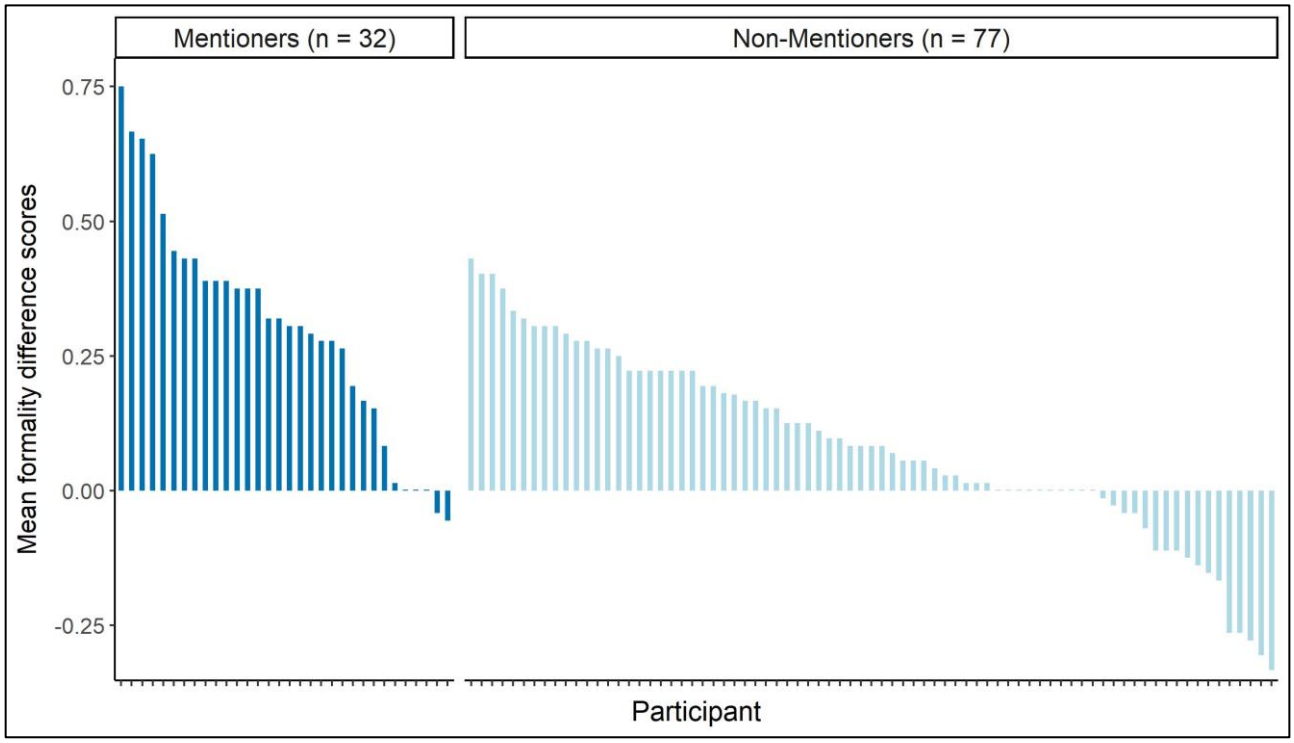


Figure 1. Mean individual difference scores: formality ratings for sentences with minus for sentences without optional liaison for Mentioners (left) and Non-Mentioners (right).

have stronger associations between acoustic cues and social categories.

The hypothesis that explicit awareness of, or attention to, socially-conditioned variation could amplify the robustness of associations between acoustic cues and social categories is not a new one. Specifically, it has been proposed in so-called dual route models of speech perception [19]; these models view interaction between social and linguistic information as an integral part of language processing. Though intuitive, it has as far as we know not been tested directly. This hypothesis could be tested in future research, especially with a wider range of methods for assessing listener awareness, including EEG [20].

Finally, the results of the current study highlight the importance of probing individual awareness of socially-indexed variation in language processing experiments. Going a step further, it would be interesting to investigate which factors account for individual differences observed in noticing socially-indexed variation during speech perception. For the case at hand, given that optional liaison is linked with political, legal, academic and business language, one may wonder if level of education plays a role. Our sample was skewed in that two thirds of participants had at least a college degree, and several of the remaining participants were college students. For what it's worth, the percentage of Mentioners did not differ between those who had a college degree and those who did not ($t < 1$). More generally, cognitive factors such as working memory and social factors such as autistic-like traits have been shown to influence sociolinguistic perception to some degree [21-22]. Whether factors such as these correlate with noticing is a promising avenue for future inquiry.

APPENDIX

A full list of sentence stimuli, each with one optional liaison environment (bolded).

1. Les népalais son**t‿h**abitués au froid.
2. Les légumes sont vraimen**t‿e**xcellents pour la santé.
3. Les bars on**t‿é**té fermés pendant longtemps.
4. Les livre**s‿a**nciens sont vendus en ligne.
5. La viande n'est pa**s‿a**u menu tous les jours.
6. Les canicules seron**t‿a**ssez fréquentes.
7. Les enfants peuven**t‿a**pprendre plusieurs langues.
8. Les américains sont vraimen**t‿a**llés sur la lune.
9. La statue vivante était parfaitemen**t‿i**mmobile.
10. Plusieurs pays veulen**t‿o**rganiser la prochaine Coupe du Monde.
11. Les couteaux suisses sont très souven**t‿u**tilisés au camping.
12. La victoire était totalemen**t‿i**nattendue.
13. Le Canada a deux langue**s‿o**fficielles.
14. Les koalas et les kangourous sont des animau**x‿a**ustraliens.
15. Les jeunes bébés boiven**t‿u**niquement du lait.
16. Le portugais est parlé au Portugal mai**s‿a**ussi au Brésil.
17. Les plat**s‿i**taliens sont très populaires.

6. REFERENCES

[1] Drager, K. & Kirtley, M. J. 2016. Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In: Babel, A. (ed), *Awareness and Control in Sociolinguistic Research.* Cambridge University Press, 1-10.

[2] Labov, W. *Sociolinguistic Patterns.* Philadelphia: Philadelphia University Press, 1972.

[3] Campbell-Kibler, K. 2010. Sociolinguistics and perception. *Language and Linguistics Compass* 4/6 377-389.

[4] Niedzielski, N. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1), 62-85.

[5] Drager, K. 2010. Sociophonetic variation in speech perception. *Language and Linguistics Compass* 4(7), 473-480.

[6] Schmidt, R. W. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11, 129-158.

[7] Squires, L. 2016. Processing grammatical differences: Perceiving versus noticing. In: Babel, A. (ed), *Awareness and Control in Sociolinguistic Research.* Cambridge University Press, 80-103.

[8] Adda-Decker, M., Fougeron, C., Gendrot, C., Delais-Roussarie, E. & Lamel, L. 2012. La liaison dans la parole spontanée familière : Une étude sur grand corpus. *Revue Française de Linguistique Appliquée* 17, 113-128.

[9] Hutin, M., Weng, C., Adda-Decker, M. & Lamel, L. 2022. La liaison facultative en français : étude de grands corpus combinant approche automatique relâchée et jugement perceptif. SHS Web of Conferences 138, *Congrès Mondial de Linguistique Française – CMLF 2022.*

[10] Delattre, P. 1947. La liaison en français, tendances et classification. *The French Review* 21(2), 148-157.

[11] D. Hornsby. *Norm and Ideology in Spoken French: A Sociolinguistic History of Liaison.* Cham, Switzerland: Springer, 2020.

[12] Dautricourt, R. G. French liaison: Linguistic and sociolinguistic influences on speech perception. Ph.D. dissertation, Dept. of Linguistics, The Ohio State Univ., Columbus, OH. 2010.

[13] Lambert, W. E. 1967. The social psychology of bilingualism. *Journal of Social Issues* 23, 91-109.

[14] N. Armstrong. *Social and Stylistic Variation in Spoken French.* Amsterdam: John Benjamins, 2001.

[15] Boersma, P. & Weenink, D. 2019. *Praat: Doing phonetics by computer* [computer program]. Available at: http://www.praat.org (accessed April 18, 2019).

[16] Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. 2017. LabVanced: a unified JavaScript framework for online studies. *International Conference on Computational Social Science*, Cologne.

[17] Bates, D., Maechler, M., Bolker, B., & Walker, S. 2015. Fitting linear mixed effects models using lme4. *Journal of Statistical Software* 67, 1–48.

[18] R Core Team (2020) *R: A language and environment for statistical computing* [software]. Vienna: R Foundation for Statistical Computing. Available at: http://www.R-project.org (accessed January 2021).

[19] Sumner M., Kim, S. K., King, E. & McGowan, K. B. 2014. The socially weighted encoding of spoken words: a dual-route approach to speech perception. *Frontiers in Psychology* 2014(4), Article 1015.

[20] Hanulikova, A., Van Alphen, P. M., Van Goch, M. M., & Weber, A. 2012. When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience* 24(4), 878-887.

[21] Lev-Ari, S., Dodsworth, R., Mielke, J. & Peperkamp, S. 2019. The different roles of expectations in phonetic and lexical processing. *Proceedings of INTERSPEECH*, Graz, Austria, 2305-2309.

[22] Hesson, A. & Shellgren, M. 2015. Discourse marker *like* in real time: Characterizing the time-course of sociolinguistic impression formation. *American Speech* 90(2), 154-186.

[1] A little differently, Dautricourt [12] shows that Metropolitan French listeners associate optional liaison with high socio-economic status (SES) speakers.

[2] In this sentence (Sentence 5, Appendix), deletion of *ne* is highly unlikely due to the surrounding phonological context.