



July 6-8, 2020  
**LabPhon17**



## Voicing-feature opacity of whispered French fricatives: the interplay between duration and position in voicing perception

Yohann Meynadier, No l Nguyen,  
Sophie Dufour

---

Aix Marseille Universit   
CNRS UMR7309

Laboratoire Parole et Langage, Aix-en-Provence,  
France  
Institut of Language, Communication and the  
Brain, Marseille, France

{yohann.meynadier ; noel.nguyen-trong ; sophie.dufour}  
@univ-amu.fr

This study focuses on the recognition of the voicing-feature in whispered speech. Our previous work in French (Dufour & Meynadier 2019) shows a later recognition of the feature [+voiced] probably based on a pre-lexical processing of acoustic information, other than the laryngeal vibration, extracted from the whispered signal and used in lexical access. Via an identification task, this study reveals that duration influences the perception of whispered fricatives' voicing: the longer /ʃ/ is, the more it is perceived as [-voiced]; the shorter /ʒ/ is, the more it is perceived as [+voiced]. This duration effect is modulated by the underlying voicing-feature and the position within the (non-)word. The fricative [+voiced] in word-final position shows a perceptual boundary particularly shifted towards durations much longer than the others.

DUFOUR S., MEYNADIER Y. (2019). Temporary ambiguity in whispered word recognition: a semantic priming study. *Journal of Cognitive Psychology*, 31(2): 157-174. DOI: 10.1080/20445911.2019.1573243.

Whispering creates **opacity of the voicing feature** due to the lack of vocal fold vibrations.

**Whispered [+voiced] obstruents are however recognized** beyond the chance threshold in various languages.

For French, Dufour & Meynadier (2019) have shown that **disambiguation of the [+voiced] feature requires additional processing time (+50 ms)**.

In this short period of time, the listener appears to exploit **acoustic traces of the voicing still present** in the whispered speech signal.

The **durations** of the consonants and that of the pre-consonant vowels are **robust correlates of voicing**.

In this study, the **duration** of the fricatives /ʃ/ and /ʒ/ is **manipulated in an identification task** in order to establish its role in voicing recognition in whispered speech.

## Hypotheses

The listener can extract acoustic information relative to consonant duration from whispered speech in order to perceptually identify the consonant's voicing feature.

**H1** – in agreement with the production data, we expect an effect of the consonant duration: the longer the consonant, the more it is perceived as [-voiced], and conversely, the shorter it is, the more it is identified as [+voiced].

**H2** – following Dufour & Meynadier (2019), if a pre-lexical processing of the acoustic cues operates, we therefore expect the duration effect not to differ for words and for non-words.

**H3** – the consonant position in the (non-)word influences the identification of the underlying voicing: due to the prosodic final lengthening, a consonant in final position in an isolated (non-)word should show a shift of the boundary of categorical perception of voicing towards a longer duration compared to a consonant in initial position.

## Experimental design

	words			non-words				
position	C	V	C	C	V	C		
initial	<i>char</i>	138	-	-	<i>cheur</i>	154	-	-
	<i>jarre</i>	106	-	-	<i>jeur</i>	111	-	-
final	<i>cache</i>	-	185	253	<i>queuche</i>	-	199	243
	<i>cage</i>	-	240	161	<i>queuege</i>	-	230	156

List of the whispered (non-)words used for the experiment, with the durations of the fricatives according to their voicing and position.

Whispered /ʃ/ and /ʒ/ were embedded in either initial or final position of CVC words and non-words forming ‘minimal pairs’ [Table, on the left].

The acoustic duration of the target fricatives was modified by resynthesis (Vocal Toolkit, Praat) on a continuum of 11 steps of 20 ms going from 50 to 250 ms. The differences in intensity were neutralized.

The remaining part of the stimulus (\_CV or CV\_) was made similar across stimuli in duration, intensity and frequency.

Two groups of 32 listeners performed an identification task (2AFC, e.g. *char* ‘tank’ vs. *jarre* ‘jar’) respectively on words and non-words for the 11 steps of the continuum in the two positions (in two separate blocks).

The stimuli were presented ten times in random order to each listener.

# Results

## H1 – effect of duration

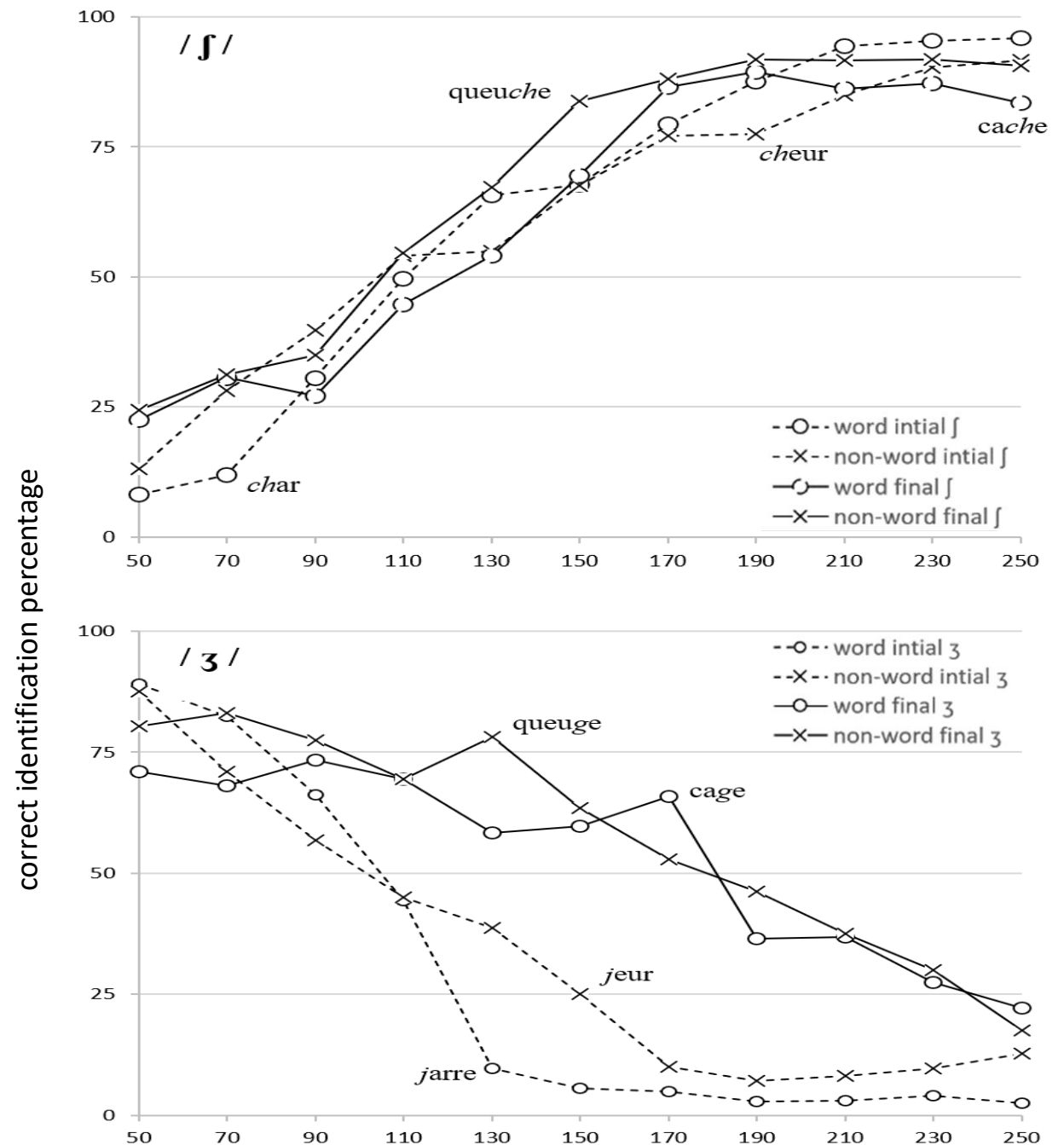
Participants' responses (1 = correct, 0 = incorrect) analyzed by a logit mixed-effects regression model.

Interaction between *duration* and *voicing* [ $\chi^2 = 5341.61$  ;  $p < .0001$ ]:

the longer /ʃ/ is, the more it is perceived [-voiced]

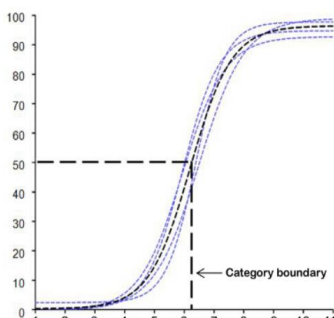
the shorter /ʒ/ is, the more it is perceived [+voiced].

This effect is observed whether the fricative is in **initial or final position of words or non-words**.

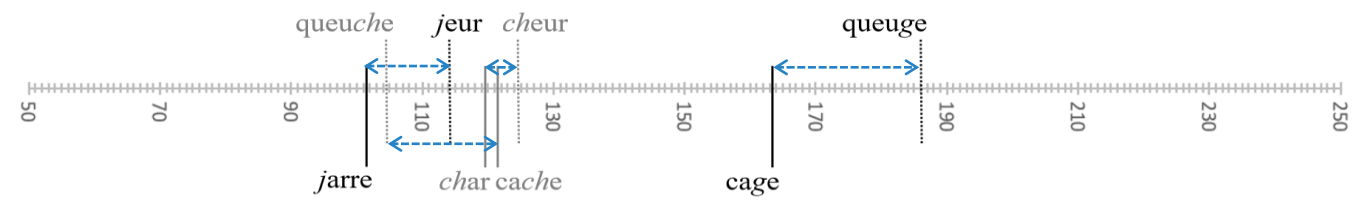
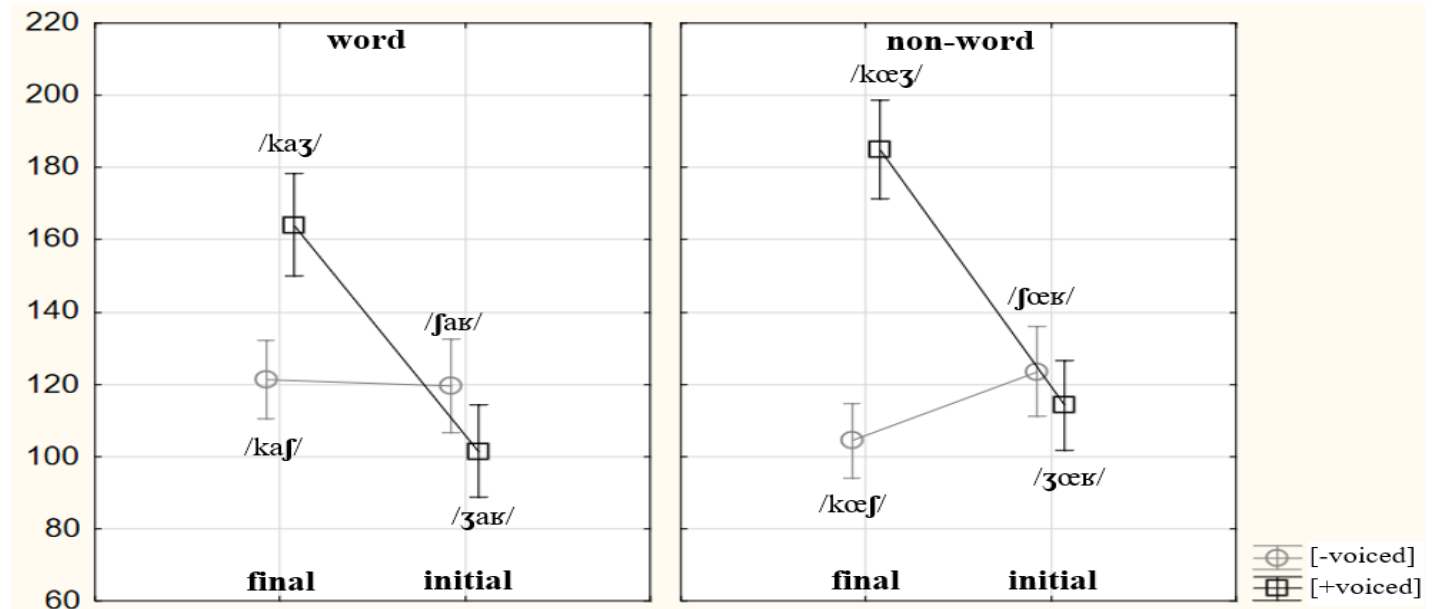


# Results

## H2 – effect of lexicality



Statistical analysis of the interactions between the duration of fricatives and voicing, position and lexicality factors was performed on the duration values of the perceptual categorical boundaries of the [±voiced] feature. The **category boundary values in the continuum** were determined by detecting the duration in ms (x) which leads to the intercept (y) between the interpolated curve and a correct response rate of 50%. The boundary values were analyzed via a three-way ANOVA including two within factors (voicing, position) and one between factor (lexicality). Post-hoc multiple pairwise comparisons (with Bonferroni correction) between the experimental conditions were also carried out.

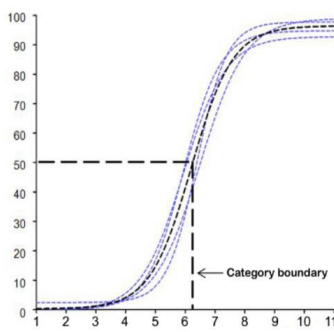


The **significant interaction between voicing and lexicality** [F (1, 54) = 12.201; p <.001] indicates that the category boundaries for the [+voiced] fricative duration have slightly higher values for non-words (queuge, jeu) than for words (cage, jar).

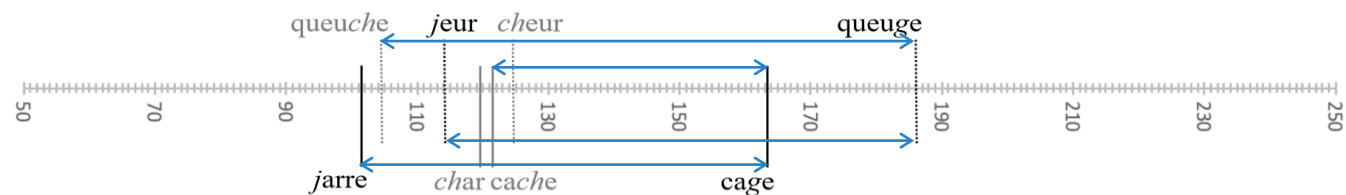
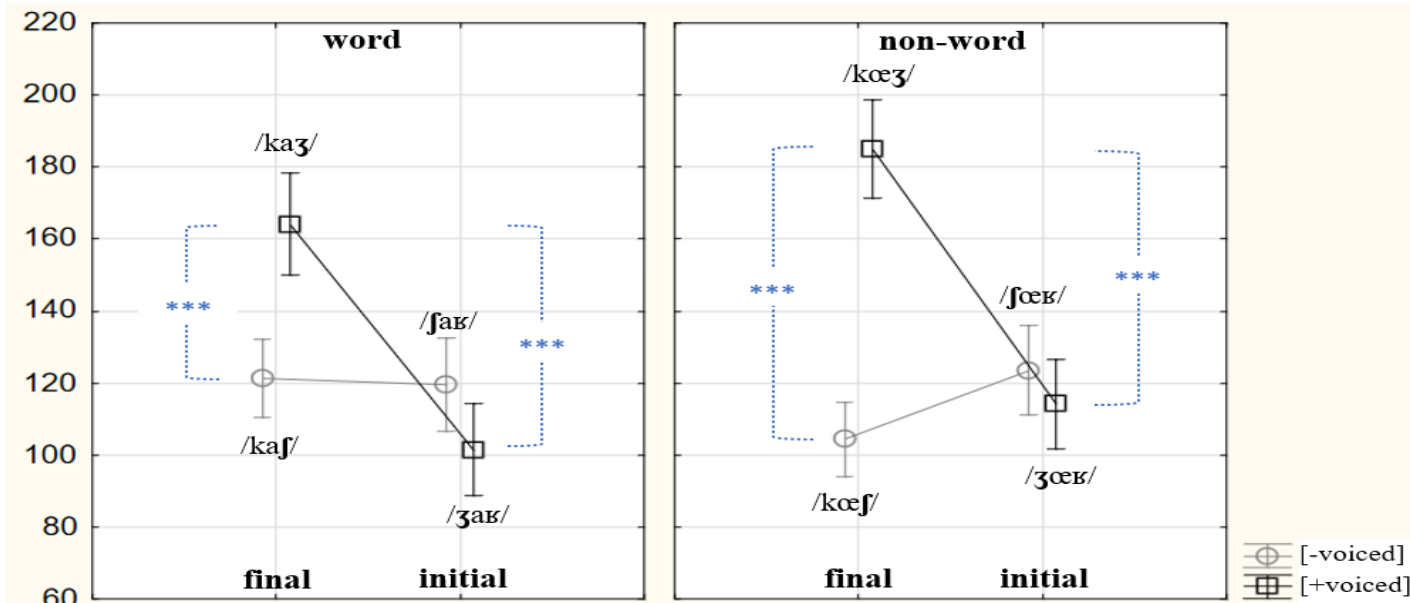
However, post-hoc tests (↔) do not indicate **any significant difference in the boundary between words and non-words.**

# Results

## H3 – effect of position



Statistical analysis of the interactions between the duration of fricatives and voicing, position and lexicality factors was performed on the duration values of the perceptual categorical boundaries of the [±voiced] feature. The **category boundary values in the continuum** were determined by detecting the duration in ms (x) which leads to the intercept (y) between the interpolated curve and a correct response rate of 50%. The boundary values were analyzed via a three-way ANOVA including two within factors (voicing, position) and one between factor (lexicality). Post-hoc multiple pairwise comparisons (with Bonferroni correction) between the experimental conditions were also carried out.



The **interaction between voicing and position is significant** [F (1, 54) = 117.60; p <.0001].

Only the **categorical boundary of the final [+voiced] fricative** (↔) is significantly higher than all the others (from 100 to 120 ms) for both words (164 ms) and non-words (186 ms).

## Discussion

**H1** – the processing of acoustic  
duration cues is confirmed

**H2** – the pre-lexical processing  
is confirmed

**H1** – The consonant duration modulates the perception of the voicing feature of whispered fricatives.

As observed in production, the subject associates the longest consonant durations with the [-voiced] feature and the shortest ones with the [+voiced] feature, regardless of the target fricative (/ʃ/ or /ʒ/) and its position as well as the lexicality of the stimuli.

**H2** – The boundaries of categorical perception of the voicing feature according to the duration of the fricative show no significant difference between words and non-words (when splitted by *voicing* and *position*).

The listeners thus seem to mainly perform a pre-lexical processing by extracting from the signal the phonetic cues and associating them with phonological units such as the feature or the phoneme, without considering the meaning or the lexicality of the phonological forms processed.



## Discussion

### H3 – the specificity of the final position is confirmed

MITTERER H., CHO T., KIM, S. (2016). How does prosody influence speech categorization? *J Phon*, 54, 68–79.  
STEFFMAN J. (2019). Phrase-final lengthening modulates listeners' perception of vowel duration as a cue to coda stop voicing. *JASA*, 154, EL560.

The effect of consonant duration is not homogeneous.

Only the [+voiced] fricative in the final position of (non-)word shows a significant shift of the category boundary of the voicing (with 40 to 80 ms) towards a consonant duration much longer compared to the (non-)word-initial [+voiced] fricative and to the [-voiced] fricative.

It remains to be explained why only the fricative [+voiced] better resists the change of duration in final position of (non-)words.

- The segments acoustically vary depending on their position and the weight of the prosodic boundary (*phonetic encoding of prosodic structure*), whose phrase-final lengthening is a specific mark [see Table in [slide #4](#)].
- The perceptual categorization of phonemes can be modulated at the contact of a prosodic boundary (Mitterer et al. 2016, Steffman 2019).
- The listener would recruit his prosodic knowledge for the recognition of segmental features allowing for a perceptual compensation for the lengthening undergone by the prosodic phrase/word-final position of the whispered [+voiced] fricative.
- The lack of effect of the final position for the whispered [-voiced] fricative could be explained by the fact that a long fricative would be interpreted by default as [-voiced]. No specific process is here needed to identify the underlying voicing.
- The whispered [+voiced] fricative lengthened in prosodic final position would remain more ambiguous. Voicing disambiguation would thus imply the recourse to prosodic information to operate at a perceptual normalization of the contextual duration for the whispered [+voiced] fricative.

## Conclusion

This study reveals a **specific behavior** of the **[+voiced]** feature and of the final prosodic position in the recognition process of the voicing feature of whispered phonemes.

The recognition of the voicing feature of whispered fricatives would thus imply a **bottom-up process** of **phonetic-detail extraction** and a **top-down process** of **prosodic compensation** of segmental durations.

The *prosodic hypothesis* is however still **speculative** at this stage.