



HAL
open science

On the perception of voicing in whisper: a cross-modal semantic priming study

Yohann Meynadier, Sophie Dufour

► **To cite this version:**

Yohann Meynadier, Sophie Dufour. On the perception of voicing in whisper: a cross-modal semantic priming study. ISP Workshop: Abstraction, Diversity, and Speech Dynamics, 2017, Herrsching am Ammersee, Germany. , ISP Workshop: Abstraction, Diversity, and Speech Dynamics, 2017, 10.13140/RG.2.2.15169.76641 . hal-01587910

HAL Id: hal-01587910

<https://hal.science/hal-01587910>

Submitted on 30 Sep 2017

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.

On the perception of voicing in whisper: A cross-modal semantic priming study

Yohann Meynadier and Sophie Dufour

Université Aix-Marseille & CNRS
Laboratoire Parole et Langage
Aix-en-Provence, France

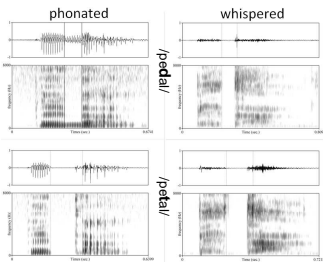


This work addresses the question of the one-line pre-lexical processing and the role of the fine phonetic details in the recognition of the voicing feature in French whispered words. As produced without vocal fold vibration, the [+voice] obstruents are not immediately recognised. But if few more time is left to the listener, they are well recognized. It suggests that this time may be used by the listener to extract from the acoustical signal the phonetic details necessary to reconstruct voicing during word recognition process.

Background & Questioning

Whisper: voiceless speech

- Open glottis > no vocal f. vibration
- ambiguity : [+voice] phonemes becomes voiceless
- not a linguistic or phonological process (vs. assimilation rules)
- a communication requirement to reduce the distance perceptibility



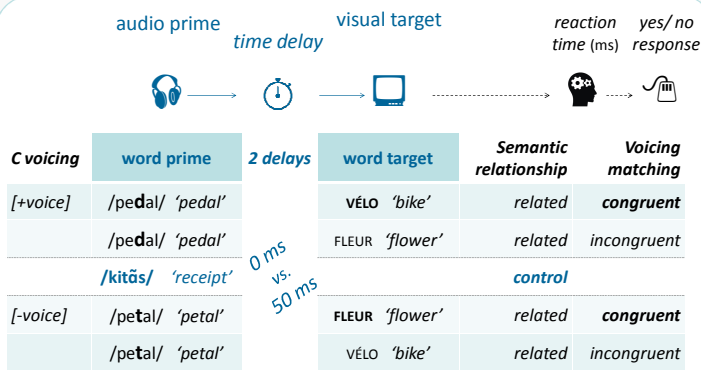
Phonetic traces of voicing

- Production: C or V durations, intra-oral pressure, glottal opening...
- Perception: few studies, difficult comparisons, off-line tasks
 - > variable levels of recognition, often better than chance
 - > a bias towards [-voice].

Phonetic details in word recognition and lexical representation

- Real-time process of phonetic details in spoken word recognition
 - > on-line processing with time pressure
- Granularity of the phonological representation of word
 - > exemplar vs. abstract

Cross-modal semantic priming experiments



Lexical task: "Is the letter sequence on the screen a French word?"

Phonated vs. whispered speech (as at the ear of a close listener)

Material and subjects

6 minimal contrasts of voicing : /p t k f s j/ vs. /b d g v z ʒ/

Primes 20 minimal pairs (*pedale-pétale*)

Targets the most frequent semantical associate of the prime > 20% from a free association task by 38 subjects (*pedale-VELO; pétale-FLEUR*)

Controls prime not semantically related to the target (*quittance-VELO/FLEUR*)

Non-word Targets = 50% (for the task)
 • 40%: pseudo semantically unrelated ('*bottle*' *bouteille-FUDUR* from 'futur')
 • 10%: pseudo semantically related ('*lettuce*' *laitue-VALADE* from *salade*)

Related prime-target stimuli = 20% (tested items)

17 French listeners per test list (N=408)
 8 test lists in whisper + 4 in phonated
 x2 prime-target delays

> each prime or target presented only once to the same listener

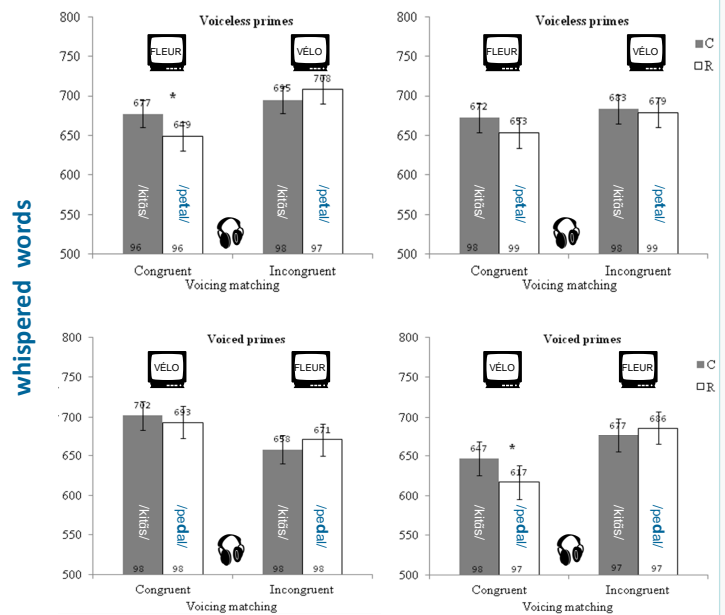
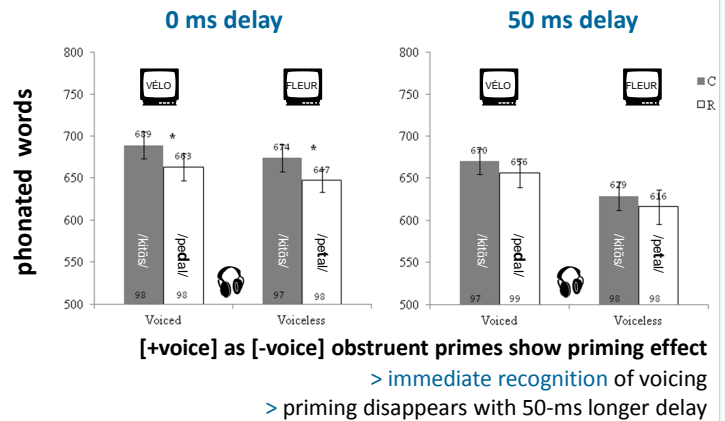
| Neutralised properties of the primes and targets | Lexical Frequency (per/mille) | Number of syllables | Number of phonemes/graphemes | Auditory uniqueness point | % of association |
|--|-------------------------------|---------------------|------------------------------|---------------------------|------------------|
| | | | | | |
| Target words (VELO) | 44 | 1.75 | 5.55 | - | 41.22 |
| Related primes (/pedal/) | 21 | 2.20 | 4.35 | 4.95 | 41.22 |
| Control primes (/kitã/) | 16 | 2.20 | 4.35 | 4.85 | - |
| Voiceless consonants | | | | | |
| Target words (FLEUR) | 106 | 1.80 | 5.90 | - | 40.30 |
| Related primes (/peta/) | 14 | 2.20 | 4.35 | 4.80 | 40.30 |
| Control primes (/kitã/) | 16 | 2.20 | 4.35 | 4.85 | - |

| Significant differences acoustical duration (ms) | Consonants | | Pre-C vowels | |
|--|------------|----------|--------------|----------|
| | [-voice] | [+voice] | [-voice] | [+voice] |
| Phonated words | 141 | 92 | 114 | 145 |
| Whispered words | 149 | 88 | 138 | 170 |

- for each phonation level separately, the intensity level of all each stimulus was normalized to the intensity level averaged over all stimuli.
- Listeners were exposed to the 'natural' sound volume characterizing each phonation: no amplification of whisper.

Results

Mean Reaction Times (in ms) and Standard Errors for the control (C) and related (R) primes. Percentages of correct responses inside the bar and significant priming effects (* p < .05)



Conclusion

- For **whispered [+voice]** obstruents
 - the recognition takes time
 - but never confounded with voiceless (ambiguity)
- > the voicing processing seems to start as soon as the whispered word is heard, but **more time is needed to identify the [+voice] feature**.
- A **reconstruction process** takes place that suggests
 - not exemplar lexical representation
 - but an **extraction of phonetic details** from the signal
- > the phonetic knowledge (i.e. of C/V durations) is used in a **pre-lexical reconstruction process of more abstract units** : phonological feature or phoneme.