Recognition of "identical" words and phrases in French

Bannert, R. and Nicolas, P. and Stridfeldt, M.

journal: TMH-QPSR
volume: 37
number: 2
year: 1996
pages: 089-092

http://www.speech.kth.se/qpsr
Recognition of “identical” words and phrases in French

Robert Bannert, Pascale Nicolas and Monika Stridfeldt.
Department of Phonetics, Umeå University.

Abstract
In a second language learning context students with a Germanic L1, in general, experience larger difficulties in understanding spoken French in contrast to other Germanic languages. The explanation is sought in the large differences in coding and decoding of words in French opposed to these languages.

In an introductory study, the claim was tested that segmentally identical phrases in French result from phonological processes like ‘liaison’, ‘enchaînement’ and consonant deletion resulting in a separate syllable tier. Twenty phrases varying from one to seven syllables and minimally contrasting in pairs or triplets were presented to French and Swedish listeners. The results show variations in different aspects. In some cases, listeners recognised the original stimulus but not in others. Swedish listeners performed partly in a different way. Reasons for the divergent behaviours will be given.

Introduction
It is a general observation that in a second language learning context students with a Germanic L1 experience larger difficulties in understanding spoken French in contrast to other Germanic languages. These fundamental problems in understanding French as a foreign language were addressed by Matter (1989). Why is it the case? The obvious answer should be that the strategies for coding and decoding words are clearly different in French compared to Germanic languages like English and German. Obviously, phonetics should be able to account for the differences. The explanations are expected to be found on the phonological, articulatory, acoustic and auditory level. In a contrastive approach, Gårding (1983) showed prosodic differences between Swedish and French.

Psycholinguistic studies (Mehler et al. 1981, Cutler et al. 1986), measuring reaction time have shown that English and French listeners behave differently when syllabifying English and French word material. In their contrastive study, Cutler et al. came to the conclusion that French listeners use syllabification in their segmentation of the words presented while English listeners do not. If this conclusion is generalised to other Germanic languages with a similar complex and unclear syllable structure as English, one implication is that Swedish learners of French have to learn a new strategy for segmenting the speech flow into semantic units like words.

In order to investigate the Swedish learners’ problems, phonetic, contrastive research has started in the project “Prosodic coding and decoding of words in fluent speech: a cross-language investigation of Swedish and French (ProCoF)”*. This acoustic and auditory investigation will concentrate on two issues: (1) the syllable type (open versus closed) and the phonological processes around it, and (2) the relationship between words and phrases in speech recognition and lexical access.

One problem that learner-listeners encounter exists in identical phrases that are the results of the phonological processes ‘liaison’ and ‘enchaînement’. It is said for instance that ‘Un bonnet blanc’ (A white bonnet) and ‘Un beau nez blanc’ (A beautiful white nose) sound identical. This claim rests solely on segmental reasoning. If prosody also is taken into consideration, this claim has to be called into question. It should be noticed that this reasoning concentrates on the acoustic information conveyed in the speech signal. In normal speech communication, contextual pragmatic and semantic cues will guide the listener to the correct (or - at least - most plausible) interpretation of the alleged ambiguous speech signal.

Aim
Therefore in an introductory study, we addressed this problem: Are alleged, segmentally identical phrases really identical? Can listeners recognise the original phrase in such a pair? If the speech signal of both
members of such a pair are identical, listeners have no access to acoustic information in guiding their recognition. Therefore listeners have to guess and the recognition rate should therefore oscillate around the 50% chance level. In this case, are listeners able to detect acoustic differences between them that obviously must exist?

Test material

The corpus consists of 20 stimulus (7 pairs and 2 triplets) and 10 distractors (Table 1). Four pairs contained just single words. The rest of the corpus consisted of phrases and sentences differing in the number of syllables.

Table 1. The test stimuli. 1 - 20. Number of syllables 1 - 6 (NS). Triplets are joined together by brackets.

<table>
<thead>
<tr>
<th>NS</th>
<th>Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>l'arme 2 larme 3 l'ange 4 lange</td>
</tr>
<tr>
<td>2</td>
<td>le son 6 leçon 7 ces jours 8 séjour</td>
</tr>
<tr>
<td>3</td>
<td>9 il est tachété 10 il est acheté 11 il est à jeter 12 Pascal pare les coups 13 Pascale parle et coud</td>
</tr>
<tr>
<td>4</td>
<td>14 il aime être au pôle 15 il aimait trop Paule 16 il est maître au pôle 17 Ca sert d’auto 18 sacerdotaux</td>
</tr>
<tr>
<td>6</td>
<td>19 Jean est ténor mais m’embêle 20 Jean est énormément bête</td>
</tr>
</tbody>
</table>

Two French speakers, a female (SA) and a male (TH), read in random order the stimuli and the distractors three times. After an auditory analysis, the second member was chosen for the test.

Listening groups

The listening test was taken by three groups: (1) the speakers themselves, (2) six French listeners in Umeå and ten in Aix-en-Provence, (3) nine Swedish learners of French at an advanced (University -C-) level. The two speakers (SA and TH) served also as listeners in order to check their capability to recognise their own productions.

Test procedure

Each stimulus was presented five times to the French listeners and three times to the Swedish listeners in random order and interspaced with the distractors. As the same stimulus is produced by each speaker, 10 and 6 responses, respectively, are given to each stimulus by each listener. Listeners were asked to write down on a response sheet the French utterance they heard as quickly as possible without further reflection. The test proper was preceded by a short training passage. The whole test was presented via loudspeakers in appropriate blocks with pauses in between.

Analysis

The total number of responses amounted to 4680. Each response is considered correct if it corresponds to the stimulus presented, although we allowed for three exceptions: (1) spelling mistakes that do not bring about confusion of meaning, (2) more than one possibility of writing (e.g. langes (plural), instead of singular - a consequence of the lack of semantical context), (3) difficulty to perceive differences between the masculine and feminine written form of Christian names.

The distribution of the correct responses over the two speakers and the three listener groups is shown in Figure 1:

![Figure 1. Percentage of correct responses obtained by each group. The speakers (SA and TH), the French listeners in Aix (FA) and in Umeå (FU) and the Swedish listeners in Umeå (SU).](image)

In all cases, we note a tendency to get better results with TH’s stimuli than with the female SA. This difference tends to be less marked for the SU group. We notice that SA has better results with TH’s voice than with her own. Thus, we expected to find more acoustic information encoded in TH’s stimuli. We obtained comparable results with the groups FA and FU. We also notice the difficulties encountered by the Swedish listeners to recognise the stimuli.
Results

Correct responses to each stimuli are shown in Figure 2 for the French listeners (the sixteen French listeners are collapsed into one group) and in Figure 3 for the Swedish listeners.

The stimuli of both speakers are treated together. The 50%-level indicates guessing which means that listeners cannot extract any acoustic cue in the signal that will guide their identification.

A first glance at both figures shows that most stimuli are recognised well above or below chance level and that the groups behave differently. For the French listeners, stimuli 1, 2, 5, 9, 14 and 17 have scores around 50%, whereas the Swedish listeners show a random recognition for stimuli 2, 5 and 8. Stimulus 19 “Jean est ténor mais m’embête”, a famous example in text books on phonological processes in French, was not recognised at all by either group. The alternative utterance heard was “Jean est énorme et m’embête”.

We also computed and analysed alternative responses obtained for the different groups. A response is considered unexpected if it does not belong to the stimulus list.

Figure 4 gives the distribution of the unexpected responses for the French listeners in Aix (FA), in Umeå (FU) and the Swedish listeners in Umeå (SU). SA female speaker, TH male speaker.

The French listeners delivered only few unexpected responses, around 8%, the Swedish listeners, however, about 25%. The male speaker’s stimuli gave rise to less unexpected responses than the female voice. This difference is absent in the Swedish listeners’ responses.
Departing from the results of the recognition task shown in Figures 2 and 3, a careful auditory and acoustic analysis of the stimuli was carried out. In some instances, we found acoustic differences that we estimate to pass the level of individual variation. In order to illustrate this aspect, Figure 5 shows the speech wave and the F0-curves of the stimuli “leçon” and “le son” spoken by the male speaker. The reason why we have chosen this example, is that we obtained a very good score of identification for this pair by the French group (around 90% of correct responses) and a random score by the Swedish group (around 50% of correct responses). As can be seen in this Figure 5, differences in the temporal structure of the utterances do not seem to be significant. However, the stimuli differ in the intensity of the syllables and in the F0. Intensity and F0 are higher in the first syllable of the word “leçon”. We think that Swedish listeners could not extract acoustical cues that permit the French to distinguish between these two stimuli.

Discussion

Our results indicate that recognition of “identical” utterances, when only the acoustic information and no contextual semantic and pragmatic cues are provided, is not random. Alleged segmentally “identical” stimuli turned out not to be identical. Listeners seem to extract some acoustic, prosodic cues that guide them in recognising the stimulus. The acoustic analysis showed that there are prosodic differences, especially intensity and F0 - not so much duration -, between the members of stimuli pairs and triplets.

This study indicates that different strategies are followed by our groups for the recognition of the stimuli. In front of a difficult task, even for the French group, we get a good insight into the behaviour of the Swedish listeners.

This study will be reported on in a more comprehensive, forthcoming paper.

* This work is supported by the Swedish Council for Research in the Humanities and Social Sciences, grant: Dnr F 176/95.

References


