UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Recognising prosodically degraded lexical material - Can word length help?

Permalink
https://escholarship.org/uc/item/2zg7b9dm

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Author
Torres, Catalina

Publication Date
2024

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Recognising prosodically degraded lexical material - Can word length help?

Catalina Torres (catalina.torres@ivs.uzh.ch)
Department of Comparative Language Science, Affolternstrasse 56
Zurich, 8050, Switzerland

Abstract

This study investigates the role of low-level acoustic cues in relation to increasing number of syllables in two word recognition experiments conducted in a remote field site. Participants are bilingual speakers of French and Drehu (Oceanic), two edge-marking languages. The two languages use grammatical gender, but differ in the number of function words used for it. In two experiments, prosodic cues were manipulated at the edges of Accentual Phrases (AP) of increasing length. APs consisted of an article and a following content word. Results show that the acoustic manipulations had a greater impact on short APs with three syllables while words in APs with more syllables could be retrieved faster in French. In Drehu, results indicate that words in longer APs are recognised later. This shows that despite similarities in the intonational phonology, listeners rely on different strategies during word recognition influenced by the grammatical make up of the language. **Keywords: prosodic cues; edge-marking; word recognition; word-length; French; Drehu; bilingualism; non-WEIRD.**

Introduction

Listeners of natural languages have at their disposal an array of strategies to help them process the incoming speech stream as it unfolds in time. To arrive at meaningful chunks such as words, listeners exploit a variety of cues to segment the speech. These range from language-specific phonological and acoustic-phonetic patterns (Cutler, 2012) to statistical regularities, or transitional probabilities (Saffran, Newport, & Aslin, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). However, in word recognition research the role of low-level acoustic cues in relation to increasing segmental material (such as syllables) represents an understudied research topic, especially in bilingual listeners. This study investigates the relationship between low-level acoustic cues and word length in spoken word recognition in listeners from a bilingual community in New Caledonia, in the South Pacific. As spoken word recognition has mostly been investigated in languages spoken in WEIRD cultures (Kapnoula, Jevtović, & Magnuson, 2024; Henrich, Heine, & Norenzayan, 2010), this study seeks to broaden this area of research by including data from a non-WEIRD community. Evidence from a number of languages has shown that acoustic correlates of prosodic cues are exploited in word segmentation. For example, duration can be used to distinguish between monosyllabic and lexically embedded words (e.g., *ham* vs. *hamster*) (Salverda, Dahan, & McQueen, 2003). Moreover, it is known that Korean listeners make use of frequent intonation patterns to identify words (Kim & Cho, 2009), and that they tune into low-level acoustic cues in this process (Tremblay, Cho, Kim, & Shin, 2019). Importantly, although duration and fundamental frequency (F0) patterns are commonly recognised as prosodic cues, the way in which they are used differs cross-linguistically (Tyler & Cutler, 2009; Tremblay, Coughlin, Bahler, & Gaillard, 2012; Tremblay, Broersma, Coughlin, & Choi, 2016; Tremblay, Kim, Shin, & Cho, 2021; Toro, Sebastián-Gallés, & Mattys, 2009; Kim, Broersma, & Cho, 2012). Further, there is evidence that faulty prosodic information carried in onset consonant clusters (voiced obstruent + liquid) affects word recognition in French (Torres & Welby, 2021). In German and French, it has been shown that the acoustic-phonetic characteristics of voiceless fricatives are systematically affected by the F0 context (Niebuhr, 2012; Welby & Niebuhr, 2016, 2019) and at least for German, it was found that listeners are sensitive to these differences (Niebuhr, 2017). This suggests that listeners are sensitive to prosodic information encoded at the segmental level. Cross-linguistic studies have focussed on monolingual listeners or second language learners (Ortega-Llebaria, Olson, & Tuninetti, 2019; Tremblay et al., 2021), and little is known on how prosodic cues are relied on in word recognition in simultaneous bilingual speakers. Since segmentation strategies are language specific, the question arises how bilingual listeners adapt to master different strategies. French and Drehu represent an interesting contrast since acoustic descriptions show that the languages share important phonetic-phonological structures while also displaying grammatical differences. At the level of intonational phonology, French and Drehu are edge-marking languages which display an F0 early rise that aligns with the content word onset (Welby, 2006; Torres, Fletcher, & Wigglesworth, 2022; Torres & Fletcher, 2020). Additionally, although French and Drehu have grammatical gender and use definite articles in prenominal position the languages differ in the number of articles they have (Moyse-Faurie, 1983).

In French, a language without lexical stress (Jun & Fougeron, 2002), lexical distinctions based on prominence marking are not stored in the mental lexicon (Michelas & Dufour, 2019) and intonation is marked at the post-lexical level. French listeners do rely on prosodic cues for word segmentation, as a number of studies have shown (Welby, 2007; Tremblay et al., 2012; Spinelli, Grimault, Meunier, & Welby, 2010; Rietveld, 1980; Banel & Bacri, 1994; Banel, Frauen-

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

felder, & Perruchet, 1998; Tyler & Cutler, 2009; Bagou & Frauenfelder, 2006). In sequences like [me.la.mɔ̃.din] when an F0 rise (early rise) started at the second syllable, listeners were more likely to perceive the sequence as two words (*mes lamondines*) but when the early rise started at the first syllable the sequence was perceived as a single content (non-)word (*melamondine*) (Welby, 2007). In contrast to English, where statistical regularities in stress pattern (predominantly word-initial stress) can be used in word segmentation by signalling potential word boundaries, the French early rise signals actual rather than potential word boundaries. A finer use of the F0 as a cue has also been identified, by which the beginning of the F0 rise at the onset of a segmentally identical pair of words (*l'affiche* vs. *la fiche*) not only helped disambiguate between them but also facilitated word recognition (Spinelli et al., 2010). Moreover, studies on word recognition have shown that a preceding article is used during word recognition (Dahan, Swingley, Tanenhaus, & Magnuson, 2000; Spinelli & Alario, 2002; Bellanger, Chevrot, & Spinelli, 2017), showing that these function words constrain the meanings that are accessed during lexical activation. However, studies examining prosodic cues to word segmentation often use short content words of one or two syllables limiting the listener to rather short input. This study seeks to examine whether listeners will make use of additional information contained in the transition probabilities between syllables. With this, the aim is to test whether listeners will be able to compensate for degraded prosodic cues to word onsets.

Relevant aspects of Lifou French

Lifou French shares most intonational properties of Hexagonal French (Welby, 2007), showing no striking differences with regard to tonal alignment to segmental landmarks (Torres et al., 2022). The initial tone L_1 is of particular interest in the present study, see Figure 1 (top). This *early elbow* has been described as a component of the phrase accent which is fairly stable, straddling the function word and content word boundary, in both Hexagonal and Lifou French (Welby, 2006; Torres et al., 2022). The early rise and elbow have been demonstrated to be used for word segmentation and represent a phonological cue to word-initial boundaries (Welby, 2007; Spinelli et al., 2010). The start point of the rise is generally associated with the left edge of a content word which explains why listeners perceive the sequence [lafiʃ] as *l'affiche* 'the poster' instead of *la fiche* 'the index card' when the rise starts earlier. Additionally, French listeners exploit final lengthening as cue to a word boundary at the right edge (Rietveld, 1980; Banel & Bacri, 1994; Banel et al., 1998; Tyler & Cutler, 2009; Bagou & Frauenfelder, 2006). As an example, lengthening can be used to discriminate between a monosyllabic word and a word-initial syllable embedded in a disyllabic lexical item (Banel & Bacri, 1994). When listeners were given ambiguous sequences of two syllables (e.g., [kɔʁ.bo]) these were interpreted as two words (*corps*, *beau* 'body, beautiful') when the first syllable was long and the second short and as one word (*corbeau* 'crow') when the first syllable was short and the second long. In other words, a iambic rhythmic pattern with a lengthened (final) syllable is associated with the demarcation of the right boundary, and in this case, with the end of a word. In French there are two forms of the definite article *la* (feminine) and *le* (masculine) and an additional form for the plural, unmarked for gender *les*. Studies using artificial language experiments have concluded that French listeners rely on word-final prominences for word segmentation (Tyler & Cutler, 2009; Bagou & Frauenfelder, 2006). However, these studies did not include function words in the artificial language, which limits the conclusions that can be made. Prominence, in these studies, is understood as a high tone or a lengthened syllable. As a result, it has been argued that French listeners do not rely on word-initial prominence when segmenting the speech stream.

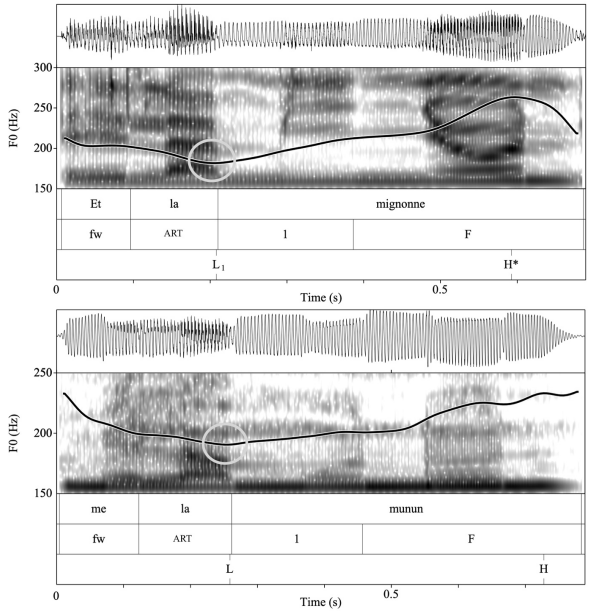

Figure 1: Illustration of two accentual phrases (AP). Top: French, *et la mignonne* 'and the cute one (feminine)' realised with a L_1H^* pattern. Bottom: Drehu, *me la munun* 'and the picot fish' realised with a LH pattern. The early elbow, circled in grey, is placed at the boundary between the article (ART) and the content word.

Relevant aspects of Drehu

Drehu is the indigenous language with the largest number of speakers in New Caledonia, and it is increasingly also spoken in the urban region of Noumea (Vernaudon, 2015; Dotte, Geneix-Rabault, & Vandeputte, 2017). The language can be seen as one of an analytic type in which particles are commonly used. As an example, there is no verbal inflexion and grammatical markers are placed in pre-verbal and pre-nominal positions (Moyse-Faurie, 1998). Relevant for this study is that there is only one form of definite article *la*, and the plural is marked with a separate particle *itre* e.g., *la itre*

koko 'the yams'. Noun phrases consisting of a definite article and a noun are mostly realised with a LH pattern that demarcates the left edge of the AP with a low tone (L) and the right edge with a high tone (H) (Torres & Fletcher, 2020). Figure 1 (bottom) shows an AP produced with the LH rising pattern.The word-initial low tone is stable despite changes in speech rate, even when words were realised with a LHLH pattern (this pattern mostly appears on long polysyllabic words). Drehu intonational phonology is similar to that of Korean (Jun, 1998), where final lengthening only occurs at Intonation Phrase boundaries (IP), but not at lower prosodic levels (Torres & Fletcher, 2022). This means that duration is not used as a cue differentiating between prosodic levels below the IP, nor does it cue word stress. However, the role of syllabic duration during word recognition has not been tested perceptually.

Research questions

This study examines reaction time during word recognition in prosodically adverse conditions. It is known that edge-marking languages rely on F0 cues to word onsets and that, at least in French, word final syllable duration is helpful during segmentation. In Exp 1 three prosodic manipulations are used to create prosodically challenging word onsets. In Exp 2 F0 cues were manipulated and additionally two manipulations are contrasted to test whether a lengthened initial or final syllable delay word recognition. To test whether increasing number of syllables and information from transition probabilities are helpful during word recognition, the experiments use APs which include a definite article and content words ranging from two to four syllables. It is hypothesised that prosodically adverse conditions at the content word onset will lead to a more delayed word recognition in short APs of three syllables. Additionally, it is hypothesised that information resulting from transition probabilities will be more informative in French than in Drehu. This is because in French there are two definite articles whereas in Drehu there is only one resulting in less informative transitions from first to second syllables. Therefore, it is predicted that Drehu words will be accessed in an incremental way as time unfolds. Additionally, this study also examines which prosodic manipulation (duration, F0, or both) has the strongest effect in delaying word recognition in French and Drehu.

Materials and Methods

Two forced choice lexical decision task experiments were conducted using auditory stimuli. In Exp 1 bilinguals were tested in French, and in Exp 2, they were tested in Drehu. Participants had to decide whether each sequence they heard contained a real word by pressing a key on a keyboard. The experiments were carried out using the software *PsychoPy*, version 3.1.5 (Peirce, 2007).

Participants

In Exp 1, 48 students (age range 14-19 years) and in Exp 2 a new group of 22 participants took part (13-48 years). Participants could perform the experiment alone in a quiet space of the school facilities, the local library, cultural centre or at their home. In line with local tradition, the participants were thanked and received a token of recognition for their time.

Materials

In Exp 1 target items consisted of each 29 two, three, and four syllables French words with a mostly CV syllable structure (the only exception is the word *pizza* which has a CVCCV structure). The items were further divided into three subgroups according to a three-way sonority distinction related to the word-initial phoneme. The groups formed consisted of (i) nasals, (ii) voiced plosives, and (iii) voiceless plosives. Note that in French the stop closure in voiced plosives has voicing (F0). The set of stimuli further included items of non-words for two, three, and four syllable words (3X30) which also had CV syllables. An additional 60 fillers were added to the target stimuli, consisting of French words with varying phonotactic structure. Stimuli consisted of a short AP containing an article followed by the target word e.g.: *la mamie* ('the granny'). A native female speaker of Hexagonal French was recorded producing all items in a carrier phrase. Target APs were excised for acoustic manipulation. The experiment included three conditions based on the three different types of manipulation carried out.

In Exp 2 a set of 45 Drehu words (19 two, 15 three, and 11 four syllable words) containing CV-syllables was constructed using a Drehu dictionary (Sam, 2009) and in consultation with a native speaker of Drehu. Additionally, 28 non-words (12 two, 8 three, and 8 four syllable words), with the same CV-structure were created. The set of stimuli also included 23 filler words with a different syllabic structure (11 two, 9 three, 3 four syllable words). A native female speaker of Drehu was recorded and the stimuli were processed in the same way as in Exp 1.

Acoustic manipulations All experimental stimuli were resynthesised. In this procedure, boundaries were moved to zero crossings before extracting the target from the carrier phrase. The acoustic manipulation was carried out using the PSOLA (pitch-synchronous overlap and add) technique in Praat (Boersma & Weenink, 2017) and scripts were used to automate the manipulations.

Exp 1: *Manipulation of fundamental frequency (RF)* the manipulation of F0 created a continuous flat pattern with no elbows or peaks. All items were produced within APs realised with a LHiL* intonational pattern, which made it possible to apply the same procedure during the manipulation. Three tonal points were identified and labelled prior to the manipulation of the stimuli: the AP initial L_1, Hi, and the final L*. The manipulation consisted of deleting the Hi tone, allowing an interpolation between L_1 and L*. *Manipulation of duration (RD)* the aim of manipulating duration was to modify syllabic duration patterns. The first syllable of each stimulus was lengthened so that the word-initial syllable was the longest in the AP. This manipulation specifically targeted

the duration of the content word-initial syllable whereas the preceding article and remaining syllables were not modified. This procedure ensured that the portion of the consonant and the vowel were increased proportionally to the original duration values. In this case, the original F0 pattern was retained and only duration was modified. *Combination of the manipulation of F0 and duration (RFD)* a third manipulation included the lengthening of the word-initial syllable as well as the flattening of the F0 curve.

Exp 2: *Duration and F0 resynthesis* the content word-initial and word-final tones were identified and then the entire contour was set to the level of the initial low tone. Then the duration of the initial syllable (RDi), and that of the final syllable (RDf) of the content word were resynthesised separately, lengthening them by a stretch factor of two. This means that the original duration of the first syllable or that of the final syllable was doubled.

Procedure Participants received a detailed oral explanation of the procedure in French. The participants were then seated in front of a computer screen and were instructed to wear and adjust a set of headphones (Audio Technica ATH M50x) that was provided for the duration of the experiment. The volume of the audio of the computer was set to a comfortable level. The experiment instructions appeared again written in French or Drehu (depending on the language tested) on the screen, additionally guiding the participants through a short test phase and allowing them to decide individually when to start the experimental trial. Participants were informed that they would hear a set of words, one by one, and that their task was to decide whether each item was or was not a real word by pressing a letter key as fast a possible. Following each response, the next stimulus was played after 1800 ms. The presentation of stimuli was randomised for each participant. All participants were exposed to all conditions in the respective experiment.

Language dominance In Exp 1: The Bilingual Language Profile (Gertken, Amengual, & Birdsong, 2014) questionnaire was collected. It showed that language dominance varied across participants (results ranged from 75 to 132 points). Taking into account the differences between participants we could identify 5 dominance groups: 11 participants were close to balanced bilinguals, 13 were dominant in Drehu, 14 were dominant in French, five showed strong dominance in Drehu and six strong dominance in French. For Exp 2 it was not possible to collect this data due to the lackof access to an internet connection which could only be guranteed for data collected in the school facilities.

Data analyses

Only correct answers were used in statistical analyses. Values with a RT greater than three times the standard deviation were discarded as well as values that were further away than 3 standard deviations from the speakers mean. Responses given after 5000 ms were considered unrealistic and discarded. Exp 1: No responses were recorded in the first 200 ms. Since the duration of stimuli varied depending on the manipulation carried out, and access to information on segmental content was delayed for stimuli with a lengthened initial syllable, the RT was calculated subtracting the duration of the stimuli and these values were used for further analyses. Further outliers were found examining the estimated densities for the distributions of reaction times and verifying the goodness of fit of the log RT, using density and Q-Q plots (Baayen & Milin, 2010). Based on this data inspection, additional data points were discarded. In total 9% of the data were removed. To control for frequency effects that could affect RTs *Lexique* was used, an online French lexical database (lexique.org), which provides frequency scores based on written texts and film subtitles (New, Pallier, Brysbaert, & Ferrand, 2004; New, Brysbaert, Veronis, & Pallier, 2007). Four words for which no frequency score could be obtained (when produced as nouns) were discarded from the analyses (3 four and 1 three syllable words[1]). The frequency score for texts and films were tested in the statistical analyses since we could not *a priori* discard differences related to exposure to either written texts or television. In Exp 2: No correct responses were recorded in the initial 800 ms. A total of 2.6% of the data was discarded from the statistical analyses. Following (Saffran et al., 1996) transition probabilities (TP) were calculated for the stimuli of each experiment. Syllable pairs were created and then the transition probability for the first to second and second to third syllables of the AP was calculated. This means that articles, as well a fillers and non-word syllables were included in the calculation of the probabilities.

Syllables/AP	**Long initial**	**Flat F0**	**Combination**
3	557 (374)	563 (371)	574 (368)
4	447 (401)	443 (349)	474 (344)
5	374 (460)	413 (484)	409 (436)

Table 1: Median RT in milliseconds, excluding outliers, subtracting duration time of stimulus. Standard deviations are given in parentheses.

Statistical analyses This study uses Bayesian mixed-effects regression analyses with the R package brms (Bürkner, 2018) and reports the median estimate from the posterior and the "probability of direction" computed with bayestestR (Makowski, Ben-Shachar, & Lüdecke, 2019). This metric indicates the percentage of the posterior with a particular sign, ranging between 50 (posterior centered directly on zero) and 100 (posterior entirely excluding zero). The measure helps quantify with what certainty an effect shows a particular directionality. For that pd>97 is taken to be robust evidence for an effect, though values approaching this threshold are considered to provide weaker evidence for

[1]The APs excluded were: *le décapité* 'the decapitated one', *le deshérité* 'the disinherited one', *le guillotiné* 'the one with his head cut off', *le nominé* 'the nominated one'.

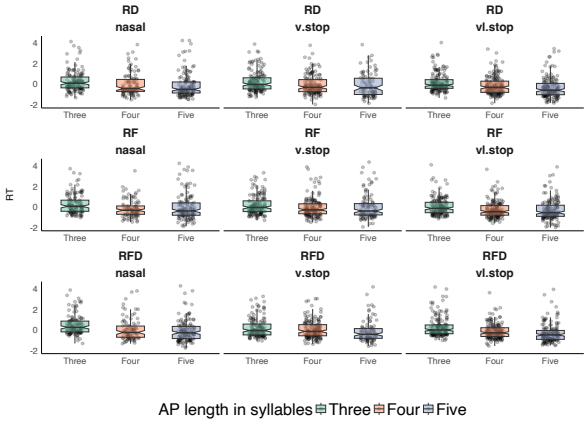

Figure 2: Scaled RT measurements in Exp 1. Content words starting with one of three segment types (nasal, voiced stop, voiceless stop) according to three manipulations (RD, RF, RFD) in polysyllabic APs (three, four, five syllables).

effect existence (e.g., if pd=92 we can be 92% sure that an effect exists with a particular directionality). Weakly informative priors are used and models were run with four chains, an adapt-delta value of 0.99, and 4,000 iterations per chain, and a burn-in period of 1,000 iterations.

Results

Experiment I

French stimuli were correctly identified in 93% of instances. Table 1 summarises RTs for correct responses given according to the three manipulations used in our experiment. Figure 2 shows RT values according to consonant at the onset of the content word and the three manipulations carried out. The scaled values of RT measurements were fitted to a linear mixed effects model to investigate the experimental manipulations (N=3613). The model had NUMBER OF SYLLABLES (three, four, five), TP FIRST TO SECOND, TP SECOND TO THIRD (continuous), MANIPULATION (RD, RF, RFD), INITIAL SEGMENT TYPE (nasal, voiced stop, voiceless stop), WORD FREQUENCY (continuous), AGE (15,16,17,18,19), and LANGUAGE DOMINANCE (continuous) as fixed factors together with participant and word as random factors, and additional random slopes for manipulation by participant and slopes for word frequency by word.

The results of the first model confirm the hypothesis that words in short APs are more strongly delayed during word recognition in French. With AP five syllables as the reference level, there is robust evidence that APs with two syllables are correctly identified later (β = 0.40, pd = 99.6%). The influence of prosodic adverse conditions rapidly slows down and there is no compelling evidence for words in APs with four syllables to be recognised with a similar delay (β = 0.16, pd = 85.9%). Results for transition probabilities show there is no compelling evidence for first to second (β = 0.19, pd = 65%) or second to third (β = 0.08, pd = 76%) to be strong predictors of reaction time. Thus, the model shows that number of syllables per AP is the strongest predictor and that in French longer words are accessed faster than short words when in prosodically adverse conditions. Further, there is no compelling evidence for either of the three prosodic manipulations having a stronger effect (RF β = 0.07, pd = 76%). However, there is a weak effect of initial segment showing that words with voiceless stops at the onset are accessed faster than words with nasal onsets (β = -0.26, pd = 96%). Post hoc pairwise comparisons confirm that words with nasal onsets were accessed later than those with voiceless onsets in APs with four (β = 0.29, pd = 97%) and five (β = 0.24, pd = 96%) syllables. There is no compelling evidence for an effect of age or language dominance, but a weak effect for word frequency (β = 0.00189, pd = 96%).

Experiment II

Results of the lexical decision task showed that participants correctly identified whether a stimulus was a real word in 84% of the cases. When exposed to real Drehu words, participants gave correct answers in 82% of instances, and when confronted with non-words in 86% of instances. Only correct responses of real Drehu words were included in the statistical analyses. The left columns in Table 2 provide the median RT responses in milliseconds for correct answers, measured from the onset of the second syllable. Columns to the right provide the mean duration of stimuli, measured from the onset of the second syllable. In the latter, the mean duration is provided as word duration corresponds to a more symmetrical distribution. RT for responses increase in parallel to the remaining word duration. However, a positive correlation between RT and the remaining time in the stimulus word is very weak (R =0.11, t = 3.0667, df = 808, *p*-value = 0.002).

	RT by condition		Mean word duration	
Syll/AP	**Initial**	**Final**	**Initial**	**Final**
3	1055 (522)	1122 (419)	309 (38)	463 (57)
4	1047 (424)	1044 (436)	576 (72)	748 (98)
5	1286 (651)	1275 (496)	708 (73)	851 (85)

Table 2: Median RT in milliseconds for correct responses, without outliers. Standard deviations are given in parentheses. Mean word duration of stimuli from onset of the second syllable, according to number of syllables and manipulation.

Figure 3 shows the scaled RT responses for APs with three, four, and five syllables. The graph indicates that RTs were slower for APs with five syllables. Note that there is no clear difference between words with a lengthened initial or final syllable. The scaled values of RT measurements were fitted in a linear mixed effects model to investigate the effects of the manipulation of duration on APs of different length (N=810). The model had NUMBER OF SYLLABLES (three, four, five), TP FIRST TO SECOND, TP SECOND TO THIRD (continuous), MANIPULATION (RDi, RDf) as fixed factors, the random factors participant and word, and also included by manipulation

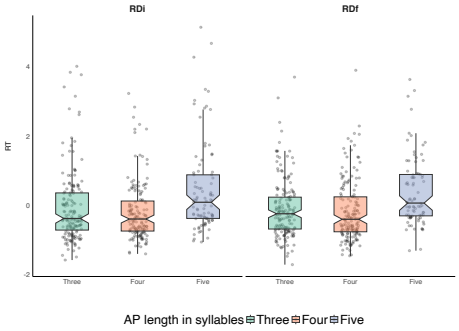

Figure 3: Scaled RT measurements in Exp 2. According to two manipulations (RDi, RDf) in polysyllabic APs (three, four, five syllables).

random intercepts for participant and word. The results of the second model confirm our hypothesis that words in short APs of four (β = -0.57, pd = 99.9%) and three (β = -0.45, pd = 99.6%) syllables are recognised faster than those in APs with five syllables. This result is also indicative that words in APs with three syllables are accessed with a relative delay compared to APs with four syllables. However, a post-hoc pairwise comparison of APs in RDi does not provide compelling evidence for this observation (β = 0.12, pd = 79.5%). Finally, no compelling evidence is found for the factors TP first to second, second to third or manipulation of first or final syllables to influence word recognition.

Discussion and conclusion

This study examines word recognition in two typologically close languages at the prosodic level but with different grammatical gender properties. The study tests the effects of word length on word recognition while prosodically degrading the edges of accentual phrases. Results show that long words can be recovered faster in French while this is not the case in Drehu. Thus, listeners exploit different strategies during word recognition. These strategies are influenced by differences in the number of definite article(s) which constrain potential word candidates in French but not in Drehu. The evidence also indicates that listeners are attuned to fine-grained language-specific phonological and acoustic cues involved in lexical recognition. Exp 1 investigated word recognition in French words of different lengths while acoustic cues at the left-edge were manipulated. Findings confirm the relevance of left-edge boundary cues to word recognition in French. Crucially, these findings show that acoustic cues are more relevant in short than in long words. This indicates that an increment in segmental material aids the listener in recognising a word, suggesting that prosodic cues can be more easily disregarded. At the same time, in absence of additional segmental material the missing acoustic cues negatively influence word recognition. It was found that content words with onsets with nasal consonants were recognised later, suggesting that the prosodic manipulations more strongly affected words with nasal onsets. This is in line with previous findings on French (Welby, 2007) and confirms the role of the F0 early rise during word recognition. Although there is no compelling evidence for the role of transition probabilities, it was fund that word length was the strongest predictor of RT in French. Taken together, this could indicate that listeners need to rely more on low-level acoustic differences when presented with short speech material and can otherwise rely on the incrementally incoming segmental material. These results confirm that although prosodic cues aid word recognition, they are part of a set of cues relied upon when listeners are performing a lexical decision task. Thus, it is found that during word recognition listeners rely on both, low-level acoustic cues as well as probabilistic knowledge that becomes available as segmental information unfolds in time. Importantly, the relevance of cues to word recognition differs with the incremental nature of language. In shorter speech sequences listeners need to rely more heavily on acoustic cues. In longer speech material the incoming segmental information can help compensate for the missing information of acoustic cues. In Exp 2, based on observations on other edge-marking languages with a similar structure to that of Drehu, it was decided that F0 cues should be eliminated. Additionally, this experiment tested whether word-initial or -final syllabic lengthening has a more negative impact on word recognition in Drehu. Our results indicate that neither the lengthening of the word-initial nor that of the word-final syllable led to consistently slower word recognition of Drehu words. We find that syllabic duration patterns are not a phonological cue facilitating word recognition. This confirms that durational cues in Drehu don't play an important role below the level of the IP, in line with previous findings (Torres & Fletcher, 2022). Interestingly, words of two and three syllables were identified at roughly the same speed. Contrary to French, in Drehu there is only one form of the definite article. Thus, while segmental information in the French articles can be used together with the incoming segmental material to retrieve potential words (Bellanger et al., 2017), the same cannot be done in Drehu. This could mean that words can only be confidently recognised as sufficient segmental material is present. The data presented in this study indicates that this threshold is reached after two content word syllables. However, more research is needed to better understand how Drehu listeners access lexical material and what role word frequency plays. Finally, this study shows differences during word recognition in the two languages of a bilingual community. Although French and Drehu are edge-marking languages sharing intonational characteristics, strategies in word recognition are not identical and word length is more helpful in French than in Drehu. This study contributes to our understanding of spoken word recognition by including data from a non-WEIRD community (Henrich et al., 2010), this emphasises the importance of carrying out experimental studies in the field which help broaden our understanding of language processing and the bilingual mind more broadly.

Acknowledgements

Special thank goes to the participants in Lifou who made this study possible. Thanks to Lizie Wea for support creating the Drehu word list and being recorded for the Drehu stimuli. Thanks to Véronique Duché for producing the French words.

References

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28.

Bagou, O., & Frauenfelder, U. H. (2006). Stratégie de segmentation prosodique: rôle des proéminences initiales et finales dans l'acquisition d'une langue artificielle. In *Acte des XXVIèmes Journées d'étude de la Parole. dinard, france.*

Banel, M.-H., & Bacri, N. (1994). On metrical patterns and lexical parsing in French. *Speech Communication*, *15*(1-2), 115–126.

Banel, M.-H., Frauenfelder, U. H., & Perruchet, P. (1998). Contribution des indices métriques à l'apprentissage d'un langage artificiel. *XXIIèmes Journées d'Etude sur la Parole*.

Bellanger, C., Chevrot, J.-P., & Spinelli, E. (2017). Gender context effects in noun recognition: grammatical cues or co-occurrence effects? *Language, Cognition and Neuroscience*, *32*(9), 1204–1209.

Boersma, P., & Weenink, D. (2017). *Praat: doing phonetics by computer (version 6.0.26)[computer program].*

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395–411. Retrieved from https://doi.org/10.32614/RJ-2018-017 doi: 10.32614/RJ-2018-017

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. MIT Press.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of memory and Language*, *42*(4), 465–480.

Dotte, A.-L., Geneix-Rabault, S., & Vandeputte, L. (2017). Nouméa at the Crossroads of New Caledonian Multilingualism. *Amerasia Journal*, *43*(1), 13–32.

Gertken, L. M., Amengual, M., & Birdsong, D. (2014). 11. assessing Language Dominance with the Bilingual Language Profile. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring l2 proficiency: Perspectives from sla* (pp. 208–225). Multilingual Matters. Retrieved from https://doi.org/10.21832/9781783092291-014 doi: doi:10.21832/9781783092291-014

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and brain sciences*, *33*(2-3), 61–83.

Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, *15*(2), 189–226.

Jun, S.-A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, *14*(1), 147–172.

Kapnoula, E. C., Jevtović, M., & Magnuson, J. S. (2024). Spoken Word Recognition: A Focus on Plasticity. *Annual Review of Linguistics*, *10*, 233–256.

Kim, S., Broersma, M., & Cho, T. (2012). The use of prosodic cues in learning new words in an unfamiliar language. *Studies in Second Language Acquisition*, *34*(3), 415–444.

Kim, S., & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *The Journal of the Acoustical Society of America*, *125*(5), 3373–3386.

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, *4*(40), 1541.

Michelas, A., & Dufour, S. (2019). Are Prosodic Variants Stored in the French Mental Lexicon? *Experimental Psychology*, *6*(66), 393–401. doi: 10.1027/1618-3169/a000462

Moyse-Faurie, C. (1983). *Le drehu, langue de Lifou (Iles Loyauté). Phonologie, morphologie, syntaxe*. Langues et Cultures du Pacifique Ivry.

Moyse-Faurie, C. (1998). Relations actancielles et aspect en drehu et en xârâcùù. *Actances*, *9*, 135–145.

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*(4), 661–677.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524.

Niebuhr, O. (2012). At the edge of intonation: The interplay of utterance-final F0 movements and voiceless fricative sounds. *Phonetica*, *69*(1-2), 7–27.

Niebuhr, O. (2017). On the perception of "segmental intonation": F0 context effects on sibilant identification in German. *EURASIP Journal on Audio, Speech, and Music Processing*, *2017*(1), 19.

Ortega-Llebaria, M., Olson, D. J., & Tuninetti, A. (2019). Explaining cross-language asymmetries in prosodic processing: The cue-driven window length hypothesis. *Language and Speech*, *62*(4), 701–736.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13.

Rietveld, A. C. M. (1980). Word boundaries in the French language. *Language and Speech*, *23*(3), 289–296.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*(2), 101–105.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, *90*(1), 51–89.

Sam, L. D. (2009). *Dictionnaire drehu-français*. Centre de Documentation Pédagogique de Nouvelle Calédonie; revised edition (2009).

Spinelli, E., & Alario, X. (2002). Gender context effects on homophone words. *Language and Cognitive Processes*, *17*(5), 457–469.

Spinelli, E., Grimault, N., Meunier, F., & Welby, P. (2010). An intonational cue to word segmentation in phonemically identical sequences. *Attention, Perception, & Psychophysics*, *72*(3), 775–787.

Toro, J. M., Sebastián-Gallés, N., & Mattys, S. L. (2009). The role of perceptual salience during the segmentation of connected speech. *European Journal of Cognitive Psychology*, *21*(5), 786–800.

Torres, C., & Fletcher, J. (2020). The alignment of F0 tonal targets under changes in speech rate in Drehu. *The Journal of the Acoustical Society of America*, *147*(4), 2947–2958. doi: 10.1121/10.0001006

Torres, C., & Fletcher, J. (2022). Phrase-level and edge marking in drehu. *Glossa: a journal of general linguistics*, *7*(1).

Torres, C., Fletcher, J., & Wigglesworth, G. (2022). Fundamental frequency and regional variation in Lifou French. *Language and Speech*, *65*(4), 889–922.

Torres, C., & Welby, P. (2021). Segments as carriers of prosodic information in word onsets. In *Proceedings of 1st international conference on tone and intonation (tai)* (pp. 61–65).

Tremblay, A., Broersma, M., Coughlin, C. E., & Choi, J. (2016). Effects of the native language on the learning of fundamental frequency in second-language speech segmentation. *Frontiers in Psychology*, *7*, 985.

Tremblay, A., Cho, T., Kim, S., & Shin, S. (2019). Phonetic and phonological effects of tonal information in the segmentation of Korean speech: An artificial-language segmentation study. *Applied Psycholinguistics*, *40*(5), 1221–1240.

Tremblay, A., Coughlin, C. E., Bahler, C., & Gaillard, S. (2012). Differential contribution of prosodic cues in the native and non-native segmentation of French speech. *Laboratory Phonology*, *3*(2), 385–423.

Tremblay, A., Kim, S., Shin, S., & Cho, T. (2021). Re-examining the effect of phonological similarity between the native-and second-language intonational systems in second-language speech segmentation. *Bilingualism: Language and Cognition*, *24*(2), 401–413.

Tyler, M. D., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America*, *126*(1), 367–376.

Vernaudon, J. (2015). Linguistic Ideologies: Teaching Oceanic Languages in French Polynesia and New Caledonia. *The Contemporary Pacific*, *27*(2), 433–462.

Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics*, *34*(3), 343–371.

Welby, P. (2007). The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication*, *49*(1), 28–48.

Welby, P., & Niebuhr, O. (2016). The influence of F0 discontinuity on intonational cues to word segmentation: a preliminary investigation. In *Proceedings of the 8th International Conference on Speech Prosody.*

Welby, P., & Niebuhr, O. (2019). Segmental intonation information in French fricatives. In *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne.*