http://www.diva-portal.org

Postprint

This is the accepted version of a paper published in . This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Borges, R. [Year unknown!]
Rapid Automatized Picture Naming as a Proficiency Assessment for Endangered Language Contexts: Results from Wilamowice
Journal of Communication and Cultural Trends

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-387662

Rapid Automatized Picture Naming as a Proficiency Assessment for Endangered Language Contexts: Results from Wilamowice

Robert Borges

Hugo Valentin Centre, Uppsala University

Abstract

This paper discusses the use of rapid automatized picture naming (RAN) in the assessment of proficiency among new speakers of endangered languages. Despite the fact that measuring proficiency among new speakers is crucial vis-à-vis development of didactic materials and understanding language change, there are often a number of practical issues that reduce the practicality of traditional language evaluation methods. This paper investigates the potential of RAN assessments to provide a suitable indication of language proficiency by means of accuracy (ability to name pictures), speed (how quickly a verbal response is produced), and cognitive control (how well the speaker mediates cognitive load while performing the task). Results from RAN assessments administered among new speakers of Wymysorys, in concert with other data collection procedures, indicate that this type of task provides accurate insight into speakers' proficiency. Latencies in the bilingual picture naming allow accurate insight into speakers' proficiency as a function of the relative degrees of language entrenchment. However, increasing cognitive load during the assessment via speed of cue stimulus and frequently switching trial language showed no effect relative to the proficiency rank order established by naming accuracy and speed.

1 Introduction

Assessment of proficiency is a crucial aspect of the study of language acquisition. In children, language is acquired in naturalistic stages manner with regular age-determined norms for the acquisition of structures, depending on the language being acquired (Meisel 2011). In second language acquisition (SLA),

acquisition also progresses in a stage-like manner, however these are mediated by other factors, like first-language interference, age of start of learning, effort invested in learning, learning strategies, quantity and quality of input, etc. (Ellis 1997; Smith, Truscott, and Hawkins 2013). From a didactic point of view, understanding the progression of an individual's proficiency level allows the educator to determine placement in a group of similarly developed individuals, who in turn receive structured input targeted to aspects of language that are still lacking. From a research perspective, the ability to track the trajectory of individuals' proficiency development will allow for a better understanding of the relationship between language variation and change.

One approach to the still-open question in sociolinguistics regarding the precise roll of synchronic variation in diachronic language development (Léglise and Chamoreau 2013) is to study new speakers of endangered languages in terms of their acquisition of proficiency, lexical and structural variants in their speech production, and the social networks within which endangered-language practices are situated. Such contexts provide an ideal venue to investigate the development of interlanguage (Selinker 1972) and its potential role in wider language change. Since new speakers tend to occupy an active role in the community and represent larger and more influential proportion of the overall speech community than learners of non-threatened languages, the context is conducive to the observation of interlanguage features that ascend to the status of model for other language learners, replacing the "native" variety as target, thus instantiating language change. In order to make claims in this regard, it is necessary to know whether observed feature variants result from the acquisition process (i.e. a speaker compensates for incomplete or inadequate acquisition of target structures) or if variants have become part of that speaker's competence. In order to assess individual speakers' competence, it is necessary therefore to implement a measure that does not rely on normative ideas about langauge correctness. This study presents results of an initial attempt at measuring langauge proficiency via Rapid Automatized picture Naming (RAN). The study was conducted in Wilamowice (Poland) with new speakers of Wymysorys, a critically endangered West Germanic language.[1]

Language assessments usually attempt to elicit a well-rounded picture of an individual's abilities in a variety of practical situations. Two well known examples of assessments for English second language speakers, the Cambridge Proficiency test and the Test of English as a Foreign Language (TOEFL), contain similar sections: reading comprehension, with written multiple choice

[1]The study was conduced under the auspices of the project *The new speakers of Wymysorys: the reconstitution of the local language and sociolinguistic identity in Wilamowice*, with support of a postdoctoral fellowship at the Hugo Valentin Centre, Uppsala University 2017–2019.

questions or fill-in-the-blanks; a writing section, where one must demonstrate mastery of the written language in terms of grammar and vocabulary, but also structuring and supporting argumentation; listening comprehension, usually accompanied by some written form of questions; speaking, with monologues and dialogues. This strategy is typical of other English-as-a-foreign-language assessments as well as assessment among learners of other languages.

Although typical, these methods have been criticized for a variety of reasons, including cultural insensitivity, influence on pedagogics (teaching for the test), test apprehension, and so on (Gee 1989). In endangered language contexts, there is an additional set of problems with assessing proficiency using such strategies. For instance, there may be no widedly accepted standardized variety, such that an evaluation of an individual's speech production may reflect differences in language variety rather than that person's actual abilities.[2] Notions of "correctness" from a pedagogical point of view are irrelevant to the entrenchment required in the development of proficiency. Similarly, a language may not have an established orthography, may have multiple competing orthographies, or not be written at all. Finally, and especially true of the case presented in this paper, there may be insufficient human resources both to produce a listening comprehension test and to conduct on-the-spot evaluations of spoken language capabilities – quite plainly, there aren't enough proficient individuals consistently available to construct and manage a comprehensive standardized test. As a result of these issues, an alternative approach for assessing entrenchment and proficiency has been sought.

In the remainder of the paper, the procedures and results of the RAN assessment are outlined. Section 2 provides brief information about Wymysorys, its speakers, and the context that led first to endangerment of the language and its subsequent revitalization. The following section discusses the concept of new speakers of endangered languages and their potential to inform understandings of language change. RAN tests are introduced in this section and a brief outline of their usage and findings in psycholingustic experimentation is given relating to the potential usage in proficiency assessment. Based on these insights, the hypotheses tested in the study are presented. Section 4 describes the data collection protocols and operationalization of the RAN assessment used in this study. Section 5 discusses the analytical procedures applied to data generated by the RAN assessment and presents results from these analyses. Implications of these analyses are discussed in terms of the usefulness of RAN as a proficiency assessment tool in Section 6. The paper concludes with an evaluation of the hypotheses presented in Section 3 and with recommenda-

[2]Or worse, that the ascription of one variety as "standard" contributes to demotivation of learners who align with non-standard varieties as well as the further marginalization and endangerment of other varieties.

tions for further testing and application of RAN as an assessment of language proficiency.

2 Language Context

Wymysorys (also Eng. *Vilamovian*, Pol. *Język Wilamowski*, endonym: *Wymysiöeryś*) is a West Germanic language, spoken primarily in and around one town, Wilamowice, in the Silesian Voivodship (Bielsko county) of Poland (Wicherkiewicz 2003; Hammarström, Forkel, and Haspelmath 2018). The area was settled in the 13$^{\text{th}}$ century by migrants from Western Europe, most probably originating from Frisian areas around the Elbe and / or Flanders. Because the area was quite thinly populated, settlers were invited from overpopulated Germanic speaking areas, and provided financial incentives to relocate by local nobility (Barciak 2001:82–85). From that time, a unique multilingual culture developed and flourished in the area. Throughout its history, Wilamowice and its people straddled boarders, and residents of the town utilized this position by establishing wide-reaching trade networks, especially dealing in textiles and horses (Wicherkiewicz 2003:10). The townspeople lived within a system of functional polyglossia where Wymysorys was widely used in the home and private situations, Polish used for religion and education, and later (under Austrian administration) German was used for commerce and administration (Ritchie 2012; Ritchie 2016; Wicherkiewicz 2003; Neels 2012).

The vitality of Wymysorys became severely threatened following the Second World War. During the war, residents of the town were ascribed status of Category 2, "of German descent", or Category 3, "Voluntarily Germanized" on the *Deutsche Volksliste* 'German Peoples List' (Wicherkiewicz 2003). In principle this was voluntary, but in practice those who did not volunteer faced severe punishment. Despite the fact that Wilamovian people did not identify with Germany or German-ness, an idea for which there is pre-war evidence, the Red Army and post-war communist government used the Volkliste as a weapon against those who had been ascribed to the list (Wicherkiewicz 2003; Wicherkiewicz 2001). In the case of Wilamowice, this meant that the language and any culturally distinct expressions (e.g. folk costumes) were banned outright in 1945; perpetrators of language and culture faced evictions, imprisonment, exile, or execution (Wicherkiewicz 2003). As such, community members were required to hide their identities, even within extended families, in order to survive. With this, people ceased using and teaching the language to their children; intergenerational transmission was abruptly stifled.

Wicherkiewicz's ominous prediction (2000, repeated in 2003 and other works) that Wymysorys would not survive the next decade, formed part of the motivation that caused the young Tymoteusz Król (b. 1994) to begin

recording audio of the language as spoken by his grandmother and her friends, eventually amassing around 800 hours of audio recordings of elderly speakers of Wymysorys, many of whom are no longer living. Around 2007, Król and his close friend Justyna Majerska (b.1994), recognizing the damning lack of didactic materials for Wymysorys, began developing these materials based on the audio recordings and teaching the language to other children on a private basis. Some of these didactic materials were eventually published (Majerska 2014; fum Dökter, Wicherkiewicz, and fum Biöetuł 2015; Król, Majerska, and Wicherkiewicz 2016). Thanks to the continued efforts of Król and Majerska, along with subsequent institutional support from Polish universities and the European Union,[3] there are now approximately twenty five individuals who self-identify as new speakers of Wymysorys. Teaching Wymysorys continues on a private basis, though there have been intermittent instances of the language being taught as an extracurricular activity in the local elementary school, even at University of Warsaw, and active communities of practice (Wenger 1999) have developed around a local cultural heritage association, theatre group, and folk music/dance troupe.

These actions have sparked promising developments with regard to the survival of Wymysorys. In addition to the growing number of active new speakers numbers, there has been a shift in attitudes towards greater acceptance of the language within the community and in the wider society. Anxiety surrounding the use of language and local customs brought about by post-war events is easing. Nils describes the prevalence of "double identity" among older speakers (2012: 128–131), which is also strongly evident among the new speakers who participated in this study. Local activists struggle for recognition from the Polish government as a linguistic minority, but the association with the "Germanness" and the volksliste continue to be used as a tool for marginalization.[4] Those new speakers who participated in the current study report that they continue with their engagement with local language and culture regardless.

[3]Notably, the projects *Linguistic heritage of Poland* 2013–2014, financed by the National Humanities Program of the Polish Ministry of Science; *Endangered languages. Comprehensive models for research and revitalization* 2013–2016 (see Olko, Wicherkiewicz, and Borges 2016) financed by the Polish Ministry of Science, under agreement 0122/NPRH2/H12/81/2013; *Documentation of Linguistic and Cultural Hertitage of Wilamowice* 2014–ongoing, financed by the National Humanities program of the Polish Ministry of Science, and; *Engaged Humanities in Europe* 2016–2019 financed by the European Union under the Horizon 2020 program, under agreement 692199.

[4]The author had opportunity to address the Polish Parliamentary Commission for National and Ethnic Minorities as a scientific expert in 2016, and was instructed by the commission to specifically provide argumentation against the idea that Wymysorys is a "dialect of German".

3 Theoretical Context and Hypotheses

A new speaker can be defined as an individual who has learned a language with little or no exposure in the home via educational programs outside the home after a community-level shift (O'Rourke, Pujolar, and Ramallo 2015:1). As a relatively recent conceptual addition in sociolinguistics, a number of recent publications (O'Rourke and Pujolar 2013; O'Rourke and Ramallo 2013; Hornsby 2015; Nance 2015; O'Rourke, Pujolar, and Ramallo 2015; O'Rourke and Pujolar 2015; Hornsby 2016; Hornsby 2017; Nance et al. 2016; Kasstan 2017; Dołowy-Rybińska 2016; Dołowy-Rybińska 2017; Smith-Christmas, Hornsby, and Moriarty 2018) have focused on areas of shared experience, language practices, and language attitudes and ideologies, specifically in terms of language purism, authenticity and ownership.

The study of new speakers provides a number of additional unique opportunities for linguistics as a discipline which so far have not been explored. Of primary importance here is that new speakers' position and role within the endangered language community tends to be more prominent in terms of influencing norms of speech behavior than a learner of a healthy majority language. In some cases, for example, new speakers outnumber remaining "native" speakers, and in extreme cases of language endangerment, or those where the last "native" speakers have already, or will soon pass away, new speakers are, or are slated to become the speech community. These observations lead to the hypothesis, around which this work is based, namely that the study of new speaker groups will allow for *observation* of the instantiation of linguistic innovation and the spread of these innovations within the speech community in situ.

In order to address this hypothesis, the interplay of variation and language change should also be understood in terms of the causes of linguistic variation. Socially and geographically conditioned variation aside, in SLA contexts —new speakers included— a major source of linguistic variation results from the acquisition process itself. Thus, it becomes necessary to understand if variation in an individual idiolect at a given moment is a representative of that person's development, or if his/her development has plateaued as a stable idiolect. This need to understand where people are in terms of their proficiency development, such that it can be determined whether a given idiolect is stable enough to be considered an interlanguage variety, potentially pose as a model for other users, and whether common feature variants are replicated from another idiolect or interlanguage, or whether they develop independently in the acquisition process.

Proficiency development involves more than learning words and grammar, however. Proficient individuals must learn how to negotiate the resources avail-

able to them. Studies show that all languages known by a multilingual individual are active all the time, and in order to utilize a single language, items from non-desired languages must be cognitively inhibited (Kroll et al. 2015; Green and Abutalebi 2013; Hermans et al. 1998; Herdina and Jessner 2002). One potential component of a proficiency assessment, moving towards a more online measurement, is Rapid Automatized picture Naming. These tests have been arguably shown to tap into the neural mechanisms responsible for aspects of language processing (Lervåg and Hulme 2009), and can thus be utilized to examine the degree to which linguistic representations are of entrenched in an individual's executive functions.

Picture naming tasks are widely used in a variety of research areas, including psychology, psycholinguistics, bilingualism research, and speech-language pathology. The picture RAN test utilized here has its roots in the latter, when Geschwind and Fusillo (1966, quoted in Denckla and Cutting 1999) used color chips to assess language deficiencies in adult stroke survivors; these individuals were unable to name colors, despite no evidence of color blindness and being able to identify the colors via matching tests. This type of rapid naming task, where a set of colors was shown with the instruction to name them sequentially, was repeated in a number of studies for different purposes and with different stimuli (letter and number graphemes, whole words, colors, and photographs/images of familiar items. The neural circuit idea that developed through observations with stroke victims led to the idea that RAN tests could be widely administered and normative results serve as predictors of cognitive function, especially reading abilities (Denckla and Cutting 1999).

Early on, it was observed that latencies in digit naming (time from presentation of numeral stimulus to producing response) among age cohorts correlated with word recognition, and served as a predictor of reading abilities; automaticity in character recognition is a key prerequisite for higher process of word recognition and reading (Spring and Davis 1988). In a similar study, Kail et al. 1999 argue that naming latencies reflect global development; naming and reading both depend on the rapid execution of underlying cognitive processes. In other words, the access to memory and automaticity in processing are key elements of both tasks. Although studies like these have been oft reproduced, the validity of naming speed as an indicator of reading abilities remains debated, especially in terms of the details of stimuli and presentation. More recently, however, a neuro-imaging study provides evidence in support of discrete RAN tests (one stimulus at a time) with image stimuli as an indicator of reading abilities, discriminating both poor and above-average readers from average readers (Cohen et al. 2018).

Picture RAN tests have been used in a number of psycholinguistic and bilingualism studies (some notable examples: Sholl, Sankaranarayanan, and

Kroll 1995; Hoshino and Kroll 2008; Hermans et al. 1998; Gollan, Starr, and Ferreira 2015; Costa and Santesteban 2004). In these studies, latencies are also the key component. As with reading abilities, naming latencies become faster in an L2 with the increased entrenchment associated with proficiency development. L2 acquisition also affects L1 naming. Ransdell and Fischler 1987, for example, showed that bilinguals were generally slower in picture naming tasks using their L1 than monolingual peers. Gollan et al. (2005) found similar results, but also illustrated that bilinguals' latency sped up with immediate repetition of the tasks, suggesting that entrenchment can be conditioned in a relatively short time. The cost of bilingualism in terms of L1 naming latencies is mediated by proficiency in the L2, where increased proficiency in L2 is costly to global L1 processing, especially in cases of immersion and / or drastic increases in L2 exposure (Costa and Santesteban 2004; van Hell and Tanner 2012).

Naming latencies can also be mediated by other factors as well. For example, frequency effects (more frequently used words are subject to faster latencies) in word recognition tasks (Diependaele, Lemhöfer, and Brysbaert 2013; de Groot et al. 2002). Meuter (1999) points out that switching languages between trials in RAN tests leads to longer latencies in L2 trials, especially in individuals who have weaker L2s. High proficiency bilinguals are reportedly not subject to these slower latencies in either of their languages (Costa and Santesteban 2004; Gullifer, Kroll, and Dussias 2013), however, suggesting that entrenchment, or strength of linguistic representations in executive function, is the crucial factor in both naming speeds and *balanced* bilingualism.

On the basis of these insights, a number of expectations can be postulated regarding the use of picture RAN tests for assessing entrenchment of individuals' languages. Firstly, study participants' accuracy rate in naming can be used as a proxy for their vocabulary size and relative exposure to the language. This exposure is a necessary prerequisite for entrenchment. Secondly, with increased proficiency in the L2, faster latencies and lower standard deviations can be expected in the L2, while simultaneously slowing L1 latencies and increasing deviations, except in the case of balanced bilinguals, who should show no significant difference in either L1/L2 latencies or latencies of monolingual or the L1 in low proficiency L2 peers. Thirdly, an observable latency effect should be apparent for non-balanced participants with an increase in cognitive interference, i.e. (a) after switches in the trial language and (b)with added cognitive load by speeding up stimuli images.

4 Data Collection Procedures

General data collection with new speakers of Wymysorys was designed to target a set of research questions and was streamlined into several tasks. A bilingual

RAN assessment administered as one task among a set of three other tasks. Data collection proceeded according to the following protocol:

1. Sociolinguistic questionnaire
2. Vidio Elicitation of Narrative
3. RAN assessment
4. Semi-structured sociolinguistic interview

This protocol generally lasted less than an hour. All the tasks, with the exception of the interview, were administered with stimulus display software OpenSesame (Mathôt, Schreij, and Theeuwes 2012). In a separate protocol, participants also worked in pairs at a director-matcher task, wherein a scene of objects was created for a director, who was tasked to instruct another participant to reconstruct the scene with an identical set of objects. This matcher was not able to see the scene the director was reconstructing. Each pair completed the task four times, with each individual playing the role of director and matcher two times each. This task produced language data with a higher degree of speaker interaction.

All sessions were audio-video recorded. A high-definition video recorder captured video of entire meeting. Similarly, an audio field recorder captured audio signal, stored to local media for the duration of each protocol. The field recorder was also hard-lined into the stimulus display laptop, which allowed audio signal from the recorder's microphones to be captured by the stimulus display laptop within each OpenSesame experiment. For the RAN assessment, this strategy produced one audio file for each response (n=100 per speaker).

The RAN assessment itself consisted of a sequence of loops wherein image stimuli are presented following a cue that informs the speaker which language to use when producing a response. Each loop consisted of the following events:

1. Fixation dot and focus "ding" (500ms) A fixation dot appears in the center of the screen and a "ding" sound is played in order to focus the participant's attention on the screen in anticipation of the next cue and stimulus. The "ding" audio cue served a secondary purpose, to index time intervals during the assessment. In the case where writing individual files per response by the stimulus display laptop would fail, results of the assessment could be recovered from internal storage on ambient recording devices.
2. Language cue (500ms) Following the fixation dot, a cue screen presented one of two cues, either a Polish flag or the coat of arms for the City of Wilamowice. These cues indicated the language in which the participant should respond to the stimulus image. Cues were presented in a fixed, pseudo-random sequence order throughout the task. The order of cue stimuli is indicated in Appendix 1.

3. Stimulus display (300–500ms) A stimulus images, generated in random order from a pool of 100 images, is presented on the screen. The pool of images, chosen from Bank of Standardized Stimuli (Brodeur, Guérard, and Bouras 2014), consists of everyday items and items featured in didactic materials developed during revitalization efforts (fum Dökter, Wicherkiewicz, and fum Biöetuł 2015; Król, Majerska, and Wicherkiewicz 2016). A list of image stimuli is provided in Appendix 2. Stimuli display speeds were varied and followed a fixed order throughout the task. OpenSesame captures audio from the moment the stimulus image is displayed.
4. Answer screen (4000ms) The stimulus display laptop's screen is blank for four seconds, allowing the participant to respond without distraction. The audio stream continues to be captured by the stimulus display laptop for the duration of this event. The audio stream is then written to a unique file and the next loop begins.

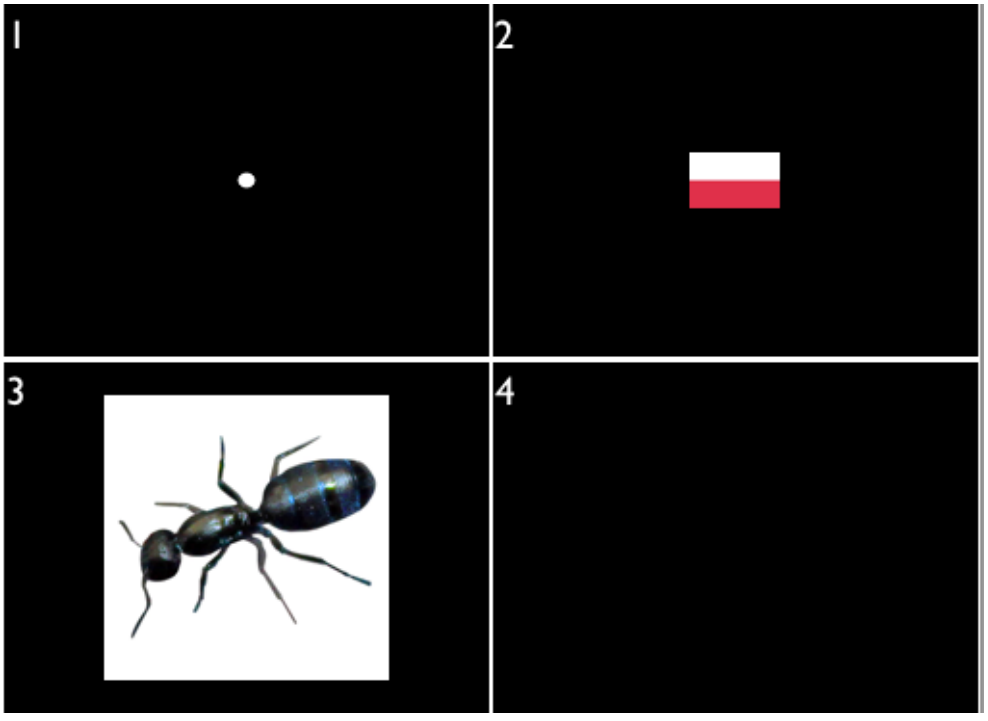

Figure 1: Example loop in the RAN sequence.

Participants were instructed to name each picture according to the cue stimulus "as quickly and accurately as possible". At the beginning of each assessment, the participant completed a practice phase, consisting of four loops,

with two Polish and two Wymysorys cues. Following the practice phase, the participant had the opportunity to ask for clarification about the instructions or the progression of the assessment. The remainder of the assessment proceeded without pause, and presented the participant with an additional 56 Wymysorys-cued and 40 Polish-cued stimuli.

Data were collected with sixteen new speakers in November and December 2017, including one reportedly balanced bilingual (the only balanced bilingual within age range of the new speakers, i.e. not elderly). The analyses presented here will focus on the responses of those speakers who were proficient and outgoing enough to complete both the narration section and RAN assessment since RAN results will be compared with proficiency measures derived from spoken data in Section 6. Seven participants met these criteria.

5 Data Analysis Procedures

Each assessment results in 100 audio responses; practice phase responses were not included in the analysis. Other responses deemed to be unacceptable were also not included; these amounted to just a handful of responses in the data. Most often post-practice phase responses were discounted as a result of the participant starting the post-practice phase of the assessment while clarification / discussion was still going on. Several other instances of interruptions, e.g. an uninvited guest entering the room, resulted in unacceptable responses.

Acceptable responses were then evaluated for accuracy (whether or not the participant named the item in the picture). Accuracy was determined rather inclusively, i.e. acceptable answers for the loop depicted in Figure 1 might be “ant”, “insect”, “bug”, or similar. Latencies were then measured for accurately named pictures. An initial attempt to mechanize this processing with a python script proved untenable. The script attempted to generate a list of response times for each response, that is, when the audio signal crosses a certain threshold, as well as a wave form and spectrogram for visual inspection against the response times. However, given the different levels of background noise, participant voice tone, and volume, even within a single assessment, it was more efficient to analyze the data manually than to continuously recalibrate the parameters of the script. Response latencies were checked manually for each response using Praat (Boersma 2002). Results were logged to a spreadsheet for further analysis. Instances of "um", "uh", and other pre-response vocalizations were not considered in the reaction time. Accuracy rates, latencies for correct responses, and probability scores calculated for each speaker’s correct responses per language are presented in Table 1. Speakers are rank ordered according to their accuracy scores.

As expected, Wymysorys latencies are significantly slower in the five speak-

Speaker	Pro	wym A	wym RT	pl A	pl RT	P	Sign.
ms9907	2	0.2	2045.5	0.93	938.0	3.13e-11	***
mz0205	2	0.23	1673.0	0.88	1356.5	0.00107	**
an0005	1	0.37	1541.5	0.85	927.0	2.11e-07	***
kz9902	6	0.43	1881.5	0.93	1022.0	8.3e-07	***
bc8612	2	0.6	1394.0	0.98	822.5	3.09e-08	***
jm9303	3	0.73	1669.5	0.83	1741.0	0.582	ø
tk9307	8	0.96	874.0	0.95	1006.5	0.0697	ø

Table 1: The columns of the Table indicate speakers' self-ascribed productive proficiency in Wymysorys (scale:0–8), "Pro"; Wymysorys and Polish accuracy rates for the RAN assessment, "wym / pl A"; Mean latencies for Wymysorys and Polish, "wym / pl RT"; The "P", probability scores are calculated based on each speaker's reaction times per language; Significance codes are Sign. codes: 0–0.001 '***', 0.001–0.01 '**', 0.01–0.05 '*' 0.05–0.1 '.', 0.1–1 'ø'.

ers with the lowest accuracy ratings in Wymysorys. The two speakers with highest accuracy rates showed no significant differences in latencies per language. Interestingly though, Polish latencies for speaker jm9303 are noticeably slower than the other speakers'. This will be taken up in Section 6. Latencies are plotted by speaker and language in Figure 2.

Contrary to expectations, however, neither stimulus display speed nor the order of stimuli produced a significant effect. Probability scores are provided in Table 2.

6 Discussion

Generally speaking, speakers' per-language latencies are close to what was expected based on the relevant literature. A general down-sloping trend is visible

Speaker	stimulus speed	order
ms9907	0.254	0.209
mz0205	0.641	0.678
an0005	0.454	0.637
kz9902	0.717	0.68
bc8612	0.54	0.765
jm9303	0.542	0.149
tk9307	0.752	0.0781

Table 2: Probability scores for the effect of stimulus speed and order in Wymysorys-cued responses.

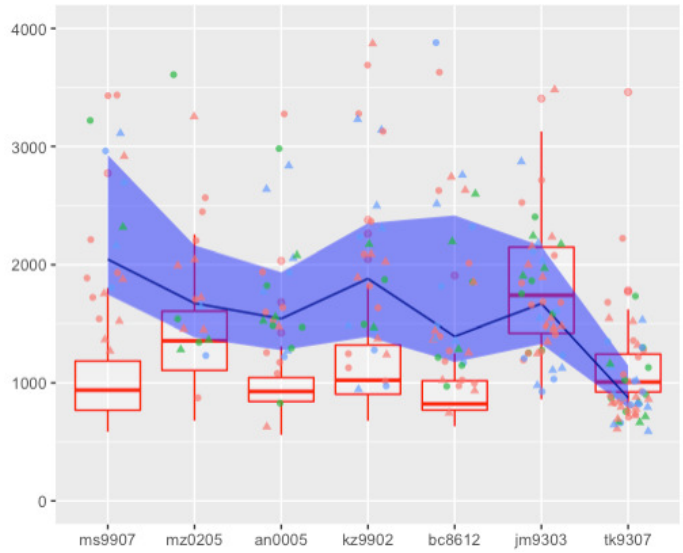

Figure 2: RAN latencies per language. Individual speakers are indexed on the x-axis; the y-axis shows the latency value in milliseconds. Red box plots illustrate speaker latencies for Polish-cued stimuli. The blue ribbon and line illustrate the standard deviation and mean, respectively, for Wymysorys-cued stimuli. A point for each Wymysorys response overlaid on the plot, further illustrating whether the previous loop was cued for the same language (= triangle) and the speed of the image stimulus (500ms=pink, 400ms=green, 300ms=blue).

in the Wymysorys-cued responses; latencies are faster with increased accuracy. The hypothesized narrowing of standard deviations is less clear, nevertheless they are certainly narrower in the balanced participant (tk9307) than the rest.

Another striking feature of the data is the difference in latencies between speakers tk9307 and jm9303. Neither speaker's response times are significantly different per language, though they are different from each other. (A full comparison of speakers' performances compared to each other can be found in Figure 3.)This can be explained in terms of entrenchment. The first speaker displays no significant differences between his Wymysorys and Polish latencies, or between most of the other participants' Polish latencies, which suggests he is able to suppress non-target representations without hesitation, supporting his self-categorisation as a balanced bilingual (Costa and Santesteban 2004; van Hell and Tanner 2012). In contrast, the second speaker, despite a very good accuracy score, and excellent abilities in Wymysorys, was significantly slower in *both* languages, suggesting a still-ongoing reorganization of linguistic representation in the cognitive processing of language. In this case, strength of

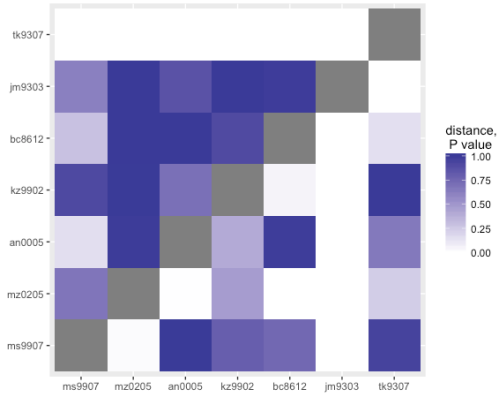

Figure 3: Distance matrix comparing speaker latencies per language. The value for generating the color gradient for each tile is the P-value generated in an ANOVA comparing latencies per speaker and language. The upper-left section displays differences among speakers in Wymysorys. The bottom-right section displays differences among speakers in Polish. Actual numerical values are given in Appendix 3.

entrenchment is such that the cost of suppressing non-target elements is costly for the L1 (Costa and Santesteban 2004; Gullifer, Kroll, and Dussias 2013).

Other potential anomalies are explained via knowledge of the speakers' engagement. The Wymysorys latencies of speaker kz9902 are relatively slow given his place in the rank order, as well as his own self-assessment. However, this speaker had been relatively unengaged with the language activities and the group of new speakers in general in the months leading up to his participation in the study. Since the automaticity from language entrenchment is something achieved by regular repetition and rehearsal, it could be postulated that his latencies were effected by atrophy. Another speaker, mz0205, produced appropriate Wymysorys latencies based on her self-assessment of proficiency and her accuracy rank order, however, her Polish latencies are significantly slower than the rest (except jm9303). Contrary to the previously discussed speaker, she is actively involved in regular Wymysorys language related activities. She also discussed some difficulties with foreign language usage —or, language block— in general during interviews. And although she didn't say much or speak very quickly during the narration task, her utterances appeared deliberate and structurally accurate. This suggests a potentially higher degree of Wymysorys entrenchment than indicated by the rank order, causing a similar suppression cost as in speaker jm9303.

One possible criticism of the study regards creating a rank order based on Wymysorys accuracy rates. It could, of course, be that participants were just

unlucky with the selection of stimuli. An attempt has been made to mediate this by choosing stimuli that represent everyday items and / or are represented in didactic materials utilized by new speakers (fum Dökter, Wicherkiewicz, and fum Biöetuł 2015; Król, Majerska, and Wicherkiewicz 2016). Nevertheless, the accuracy rank order is compared with other hypothesized measures of proficiency, such as speakers' own self assessment, speech rate, "um" rate, and lexical diversity[5] (Polinsky 2008; Irizarri van Suchtelen 2016). These measures were performed on each speaker's narration of the well known "Pear film"; results are presented in Table 3.

None of these additional measurements, purported to be indicative of proficiency, replicate the rank order derived from Wymysorys accuracy, nor do they provide a better fit to the latencies slope. While it is suspect that none of these measures replicate the same rank order, they have not yet been fully tested here, especially in regards to the minimum amount of data to achieve consistent results. Just one six-minute stimulus probably does not provide an adequate supply of data. Nevertheless, it seems prudent to continue with multiple types of assessments so as not to misrank the speaker who, for example, just says "um" a lot, or just got unlucky with the selection of stimuli, but otherwise uses the language in a highly proficient way.

A final point for discussion addresses the apparent lack latency effect af-

[5]The potential redundancy between lexical diversity and accuracy rate is recognized. Both are indications of vocabulary size and relate to entrenchment in the same way.

Speaker	Pro	wpm1	wpm2	um tok	um tim	tot. words	lex. div
ms9907	2	111.4	52.4	45	38.3s	312	0.28
an0005	1	96.6	23.9	23	15.1s	142	0.39
kz9902	6	165.2	75.8	4	3.86s	451	0.36
bc8612	2	116.8	66.7	66	57.1s	397	0.28
jm9303	3	119.3	71.6	14	7.7s	484	0.37

Table 3: Comparison of proposed proficiency measures. Speakers are organized in their Wymysorys accuracy-based rank order. "Pro" is each speaker's self assessment for Wymysorys productive proficiency; Words per minute are indicated as a calculation based on the duration of the entire stimulus film –"wpm1"– and based on the speakers' actual talking time (>0.5s silence omitted) –"wpm2". "um tok" lists the number of "um" tokens produced by each speaker, while "um tim" is the sum time of the speaker's occurrence of "um". Total number of words each speaker used in the film is indicated under "tot. words" and lexical diversity indicates total unique words divided by total number of words. N.B. despite completing the other films in the narration task, speaker mz0205 froze during this particular stimulus and is thus omitted from the table. (scale:0–8)

ter intentional stressors on cognitive processing by switching cue stimuli and varying stimulus display speeds during the RAN assessment. One can only speculate at this point about the lack of latency effect in the ordering of stimuli. The fact that switching language cues occurred throughout the test perhaps resulted in quick habituation of cognitive resources needed (cf. Schuch and Grange 2018) to mediate the tasks. This is perhaps an indication that the study design should be revised, utilizing alternative chunking patterns, for instance with longer stretches of cues in one language or the other. Speed of the stimuli was another factor intended to strain cognitive resources, thus slowing down latencies in less entrenched language varieties. The complete lack of effect may indicate that the stimuli speeds were too close together.

7 Conclusion

This paper explored the idea of using Rapid Automatized picture Naming tests as a component in assessing proficiency in Endangered language contexts. This is necessary because creation and application of traditional standardized tests in these contexts is problematic for a number of reasons, outlined earlier section. Using data collected among new speakers of Wymysorys the hypotheses put forward were tested.

First, it was hypothesized that accuracy rates in naming pictures in Wymysorys is representative of vocabulary size in the language and can be taken as proxy for general amount of exposure to the language. A satisfactory rank order of participants was created based on these accuracy rates, though this order was not recreated in either speaker self-assessment of productive proficiency or other measures implemented on spoken data, such as speech rate, "um" rate, or lexical diversity. Secondly, the data support the hypotheses that reaction times can be used as a tool for assessing proficiency as a function of language entrenchment. Lower-proficiency speakers all showed significant differences in latencies per language. The two higher-proficiency speakers showed no significant differences in latencies between languages, and while one speaker with balanced high proficiency produced no difference in latencies compared to the L1 of lower-proficiency participants, the other unbalanced, high-proficiency speaker's results illustrated the cognitive load of mediating languages with differential entrenchment. An attempt to stress the cognitive processing and thereby slow down latencies in less proficient participants via regular switching of language cues and alternating the speed of stimuli failed across the group of study participants, contrary to expectations based on previous studies. This suggests that procedural aspects of the assessment described here should be revised in future research.

In addition to these revisions, the assessment described here should undergo

further testing. At the time of this draft, a second test following the same protocols has been administered with the same cohort of Wymysorys new speakers; analysis of these tests will be presented some time in the future. However, at present Wymysorys new speakers make up an extremely small group and the assessment described here must certainly be tested on larger groups of new speakers of endangered languages and control groups of larger-language bilinguals, before a suggestion of more generalizable effectiveness could be posited with some degree of certainty. It is worthwhile to continue developing proficiency assessments that do not rely on prescriptivist notions of the language in question. Such assessments can be employed in small, endangered language contexts, where scarcity of human resources and the potential lack of standardization do not allow for more traditional proficiency assessments to be performed. Alternative assessemnts that measure proficiency as a function of langauge entrenchment, such as RAN, also bring us closer to understanding the relationship between proficiency and langauge acquisition, on the one hand, and language variation and change, on the other.

References

Barciak, Antorni (2001). *Wilamowice. Przyroda, Historia, Język, Kultura, Oraz Społeczeństwo Miasta i Gminy*. Wilamowice: Urząd Gminy.

Boersma, P. (2002). "Praat, a System for Doing Phonetics by Computer". en. In: *Glot International* 5.

Brodeur, Mathieu B., Katherine Guérard, and Maria Bouras (Sept. 2014). "Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos". en. In: *PLOS ONE* 9.9, e106953. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0106953.

Cohen, Marjolaine et al. (Feb. 2018). "Does the Relation between Rapid Automatized Naming and Reading Depend on Age or on Reading Level? A Behavioral and ERP Study". In: *Frontiers in Human Neuroscience* 12. ISSN: 1662-5161. DOI: 10.3389/fnhum.2018.00073.

Costa, Albert and Mikel Santesteban (May 2004). "Lexical Access in Bilingual Speech Production: Evidence from Language Switching in Highly Proficient Bilinguals and L2 Learners". en. In: *Journal of Memory and Language* 50.4, pp. 491–511. ISSN: 0749596X. DOI: 10.1016/j.jml.2004.02.002.

de Groot, Annette M.B. et al. (July 2002). "Lexical Decision and Word Naming in Bilinguals: Language Effects and Task Effects". en. In: *Journal of Memory and Language* 47.1, pp. 91–124. ISSN: 0749596X. DOI: 10.1006/jmla.2001.2840.

Denckla, Martha Bridge and Laurie E. Cutting (Dec. 1999). "History and Significance of Rapid Automatized Naming". en. In: *Annals of Dyslexia* 49.1, p. 29. ISSN: 1934-7243. DOI: 10.1007/s11881-999-0018-9.

Diependaele, Kevin, Kristin Lemhöfer, and Marc Brysbaert (May 2013). "The Word Frequency Effect in First- and Second-Language Word Recognition: A Lexical Entrenchment Account". en. In: *Quarterly Journal of Experimental Psychology* 66.5, pp. 843–863. ISSN: 1747-0218, 1747-0226. DOI: 10.1080/17470218.2012.720994.

Dołowy-Rybińska, Nicole (2016). "Becoming an Activist: A Self-Representation of Young European Campaigners for Minority Languages". In: *Integral Strategies for Language Revitalization*. Warsaw: University of Warsaw, pp. 405–436.

— (2017). *"Nikt za nas tego nie zrobi": praktyki językowe i kulturowe młodych aktywistów mniejszości językowych Europy*. Polish. OCLC: 1003893171. ISBN: 978-83-231-3722-1.

Ellis, Rod (1997). *Second Language Acquisition*. en. OUP Oxford. ISBN: 978-0-19-437212-1.

fum Dökter, Tiöma [T. Król], Tomasz Wicherkiewicz, and Jüśja [J. Majerska] fum Biöetuł (2015). *Ynzer Jyśta Wjytła*. Warsaw: University of Warsaw.

Gee, James Paul (Jan. 1989). "Literacy, Discourse, and Linguistics: Introduction". In: *Journal of Education* 171.1, pp. 5–17. ISSN: 0022-0574. DOI: 10.1177/002205748917100101.

Gollan, Tamar, Rosa Montoya, et al. (Oct. 2005). "Bilingualism Affects Picture Naming but Not Picture Classification". en. In: *Memory & Cognition* 33.7, pp. 1220–1234. ISSN: 0090-502X, 1532-5946. DOI: 10.3758/BF03193224.

Gollan, Tamar, Jennie Starr, and Victor Ferreira (Feb. 2015). "More than Use It or Lose It: The Number-of-Speakers Effect on Heritage Language Proficiency". en. In: *Psychonomic Bulletin & Review* 22.1, pp. 147–155. ISSN: 1069-9384, 1531-5320. DOI: 10.3758/s13423-014-0649-7.

Green, David W. and Jubin Abutalebi (Aug. 2013). "Language Control in Bilinguals: The Adaptive Control Hypothesis". In: *Journal of Cognitive Psychology (Hove, England)* 25.5, pp. 515–530. ISSN: 2044-5911. DOI: 10.1080/20445911.2013.796377.

Gullifer, Jason, Judith Kroll, and Paola Dussias (2013). "When Language Switching Has No Apparent Cost: Lexical Access in Sentence Context". In: *Frontiers in Psychology* 4. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2013.00278.

Hammarström, Harald, Robert Forkel, and Martin Haspelmath, eds. (2018). *Glottolog 3.3*. Jena: Max Planck Institute for the Science of Human History.

Herdina, Philip and Ulrike Jessner (2002). *A Dynamic Model of Multilingualism: Perspectives of Change in Psycholinguistics*. English. OCLC: 718292051. Clevedon: Multilingual Matters. ISBN: 978-1-85359-468-7 978-1-85359-467-0.

Hermans, Daan et al. (Dec. 1998). "Producing Words in a Foreign Language: Can Speakers Prevent Interference from Their First Language?" In: *Bilingualism: Language and Cognition* 1.3, pp. 213–229. ISSN: 13667289, 14691841. DOI: 10.1017/S1366728998000364.

Hornsby, Michael (2015). *Revitalizing Minority Languages: New Speakers of Breton, Yiddish and Lemko*. English. OCLC: 909320869. ISBN: 978-1-137-49879-3.

— (2016). ""Varieties of Variation in a Very Small Place" Revisited: Some Considerations from Wilamowice". In: *Integral Strategies for Language Revitalization*. Ed. by Justyna Olko, Tomasz Wicherkiewicz, and Robert Borges. Warsaw: University of Warsaw, pp. 81–90.

— (Jan. 2017). "Finding an Ideological Niche for New Speakers in a Minoritised Language Community". In: *Language, Culture and Curriculum* 30.1, pp. 91–104. ISSN: 0790-8318. DOI: 10.1080/07908318.2016.1230622.

Hoshino, Noriko and Judith Kroll (Jan. 2008). "Cognate Effects in Picture Naming: Does Cross-Language Activation Survive a Change of Script?"

en. In: *Cognition* 106.1, pp. 501–511. ISSN: 00100277. DOI: 10.1016/j.cognition.2007.02.001.

Irizarri van Suchtelen, Pablo (2016). "Spanish as a Heritage Language in the Netherlands: A Cognitive Linguistic Exploration". English. OCLC: 959743242. PhD thesis.

Kail, Robert, Lynda K. Hall, and Bradley J. Caskey (June 1999). "Processing Speed, Exposure to Print, and Naming Speed". en. In: *Applied Psycholinguistics* 20.2, pp. 303–314. ISSN: 1469-1817, 0142-7164.

Kasstan, Jonathan (Aug. 2017). "New Speakers: Challenges and Opportunities for Variationist Sociolinguistics". en. In: *Language and Linguistics Compass* 11.8, e12249. ISSN: 1749818X. DOI: 10.1111/lnc3.12249.

Król, Tymoteusz, Justyna Majerska, and Tomasz Wicherkiewicz (2016). *Hećі Pećі*. Warsaw: University of Warsaw.

Kroll, Judith et al. (2015). "Selection and Control in Bilingual Comprehension and Production". In: *The Cambridge Handbook of Bilingual Processing*. Ed. by John W. Schwieter. Cambridge: Cambridge University Press, pp. 485–507. ISBN: 978-1-107-44725-7. DOI: 10.1017/CBO9781107447257.021.

Léglise, Isabelle and Claudine Chamoreau (2013). "Variation and Change in Contact Settings". en. In: *Studies in Language Variation*. Ed. by Isabelle Léglise and Claudine Chamoreau. Vol. 12. Amsterdam: John Benjamins Publishing Company, pp. 1–20. ISBN: 978-90-272-3492-6 978-90-272-7248-5. DOI: 10.1075/silv.12.01leg.

Lervåg, Arne and Charles Hulme (Aug. 2009). "Rapid Automatized Naming (RAN) Taps a Mechanism That Places Constraints on the Development of Early Reading Fluency". eng. In: *Psychological Science* 20.8, pp. 1040–1048. ISSN: 1467-9280. DOI: 10.1111/j.1467-9280.2009.02405.x.

Majerska, Justyna (2014). *Wymysiöer Fibl*. Warsaw: University of Warsaw.

Mathôt, Sebastiaan, Daniel Schreij, and Jan Theeuwes (June 2012). "OpenSesame: An Open-Source, Graphical Experiment Builder for the Social Sciences". en. In: *Behavior Research Methods* 44.2, pp. 314–324. ISSN: 1554-3528. DOI: 10.3758/s13428-011-0168-7.

Meisel, Jürgen M. (2011). *First and Second Language Acquisition: Parallels and Differences*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press. ISBN: 978-0-521-55294-3 978-0-521-55764-1.

Meuter, Renata F.I. and Alan Allport (Jan. 1999). "Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection". en. In: *Journal of Memory and Language* 40.1, pp. 25–40. ISSN: 0749596X. DOI: 10.1006/jmla.1998.2602.

Nance, Claire (Sept. 2015). "'New' Scottish Gaelic Speakers in Glasgow: A Phonetic Study of Language Revitalisation". en. In: *Language in Society* 44.4, pp. 553–579. ISSN: 0047-4045, 1469-8013. DOI: 10.1017/S0047404515000408.

Nance, Claire et al. (Apr. 2016). "Identity, accent aim, and motivation in second language users: New Scottish Gaelic speakers' use of phonetic variation". gd. In: *Journal of Sociolinguistics* 20.2, pp. 164–191. ISSN: 1467-9841. DOI: 10.1111/josl.12173.

Neels, Rinaldo (2012). "De Nakende Taaldood van Het Wymysojer in Wilamowice, Een Germaans Taaleiland in Zuid-Polen. Een Sociolinguïstische Analyse." PhD thesis. Katholieke Universiteit Leuven.

O'Rourke, Bernadette and Joan Pujolar (2013). "From Native Speakers to "New Speakers" – Problematizing Nativeness in Language Revitalization Contexts". eng. In: *Histoire Épistémologie Langage* 35.2, pp. 47–67.

— (2015). "New Speakers and Processes of New Speakerness across Time and Space". In: *Applied Linguistics Review* 6.2, pp. 145–150. ISSN: 1868-6311. DOI: 10.1515/applirev-2015-0007.

O'Rourke, Bernadette, Joan Pujolar, and Fernando Ramallo (2015). "New Speakers of Minority Languages: The Challenging Opportunity. Foreword". In: *International Journal of the Sociology of Language* 2015.231, pp. 1–20. ISSN: 1613-3668. DOI: 10.1515/ijsl-2014-0029.

O'Rourke, Bernadette and Fernando Ramallo (June 2013). "Competing Ideologies of Linguistic Authority amongst New Speakers in Contemporary Galicia". en. In: *Language in Society* 42.3, pp. 287–305. ISSN: 0047-4045, 1469-8013. DOI: 10.1017/S0047404513000249.

Olko, Justyna, Tomasz Wicherkiewicz, and Robert Borges (2016). *Integral Strategies for Language Revitalization*. Warsaw: University of Warsaw.

Polinsky, Maria (2008). "Gender under Incomplete Acquisition: Heritage Speakers' Knowledge of Noun Categorization". In: *Heritage Language Journal* 6.1.

Ransdell, Sarah Ellen and Ira Fischler (Aug. 1987). "Memory in a Monolingual Mode: When Are Bilinguals at a Disadvantage?" en. In: *Journal of Memory and Language* 26.4, pp. 392–405. ISSN: 0749596X. DOI: 10.1016/0749-596X(87)90098-2.

Ritchie, Carlo (2012). "Some Considerations on the Origins of Wymysorys". PhD thesis. University of Sydney.

— (2016). "Language Visibility and Wymysorys". In: *Integral Strategies for Language Revitalization*. Ed. by Justyna Olko, Tomasz Wicherkiewicz, and Robert Borges. Warsaw: University of Warsaw, pp. 65–80.

Schuch, Stefanie and James A. Grange (May 2018). "Increased Cognitive Control after Task Conflict? Investigating the N-3 Effect in Task Switching". en. In: *Psychological Research*. ISSN: 0340-0727, 1430-2772. DOI: 10.1007/s00426-018-1025-4.

Selinker, Larry (1972). "INTERLANGUAGE". In: *IRAL - International Review of Applied Linguistics in Language Teaching* 10.3, pp. 209–232. ISSN: 1613-4141. DOI: 10.1515/iral.1972.10.1-4.209.

Sholl, Alexandra, Aruna Sankaranarayanan, and Judith Kroll (Jan. 1995). "Transfer Between Picture Naming and Translation: A Test of Asymmetries in Bilingual Memory". en. In: *Psychological Science* 6.1, pp. 45–49. ISSN: 0956-7976, 1467-9280. DOI: 10.1111/j.1467-9280.1995.tb00303.x.

Smith-Christmas, Cassie, Michael Hornsby, and Máiréad Moriarty (2018). *New Speakers of Minority Languages: Linguistic Ideologies and Practices*. English. OCLC: 1030362572. ISBN: 978-1-137-57557-9.

Smith, Michael Sharwood, John Truscott, and Roger Hawkins (2013). "Explaining Change in Transition Grammars". In: *The Cambridge Handbook of Second Language Acquisition*. Ed. by Julia Herschensohn and Martha Young-Scholten. Cambridge: Cambridge University Press, pp. 560–580. ISBN: 978-1-139-05172-9. DOI: 10.1017/CBO9781139051729.033.

Spring, Carl and John M. Davis (Dec. 1988). "Relations of Digit Naming Speed with Three Components of Reading". en. In: *Applied Psycholinguistics* 9.4, pp. 315–334. ISSN: 1469-1817, 0142-7164. DOI: 10.1017/S0142716400008031.

van Hell, Janet G. and Darren Tanner (Sept. 2012). "Second Language Proficiency and Cross-Language Lexical Activation". en. In: *Language Learning* 62, pp. 148–171. ISSN: 1467-9922. DOI: 10.1111/j.1467-9922.2012.00710.x.

Wenger, Etienne (Sept. 1999). *Communities of Practice: Learning, Meaning, and Identity*. en. Cambridge University Press. ISBN: 978-0-521-66363-2.

Wicherkiewicz, Tomasz (2000). "He Impact of Politics and Social Factors on the Death of a Minority Language (the Case of Wilamowicean in Poland)". In: *Developing Minority Languages The Proceedings of the Fifth International Conference on Minority Languages*. Ed. by P.W. Thomas and J. Mathias. Cardiff: Gomer, pp. 547–555.

— (2001). "Piśmiennictwo w Etnolekcie Wilamowskim". In: *Wilamowice*. Ed. by Antorni Barciak. Wilamowice: Urząd Gminy, pp. 520–538.

— (2003). *The Making of a Language, The Case of the Idiom of Wilamowice, Southern Poland*. ENGL. Reprint 2011. Berlin, Boston: De Gruyter Mouton. ISBN: 978-3-11-090540-3. DOI: 10.1515/9783110905403.

Appendix 1: Cue Stimuli order

Loop	Lang. Cue	Cue speed
1	Polish	500
2	Polish	500
3	Wymysorys	500
4	Wymysorys	500
5	Wymysorys	500
6	Wymysorys	500
7	Wymysorys	500
8	Polish	500
9	Polish	500
10	Polish	500
11	Wymysorys	500
12	Wymysorys	500
13	Polish	500
14	Wymysorys	500
15	Polish	500
16	Wymysorys	500
17	Polish	500
18	Wymysorys	400
19	Wymysorys	400
20	Polish	400
21	Polish	400
22	Wymysorys	400
23	Wymysorys	300
24	Polish	300
25	Wymysorys	300
26	Wymysorys	300
27	Polish	300
28	Wymysorys	300
29	Wymysorys	500
30	Wymysorys	500
31	Wymysorys	500
32	Polish	500
33	Polish	500
34	Polish	500
35	Wymysorys	500
36	Wymysorys	500
37	Polish	500
38	Wymysorys	500
39	Polish	500
40	Wymysorys	500
41	Polish	500
42	Wymysorys	400
43	Wymysorys	400
44	Polish	400
45	Polish	400
46	Wymysorys	400
47	Wymysorys	300
48	Polish	300
49	Wymysorys	300
50	Wymysory	300
51	Polish	300
52	Wymysorys	300
53	Wymysorys	500
54	Wymysorys	500
55	Wymysorys	500
56	Polish	500
57	Polish	500
58	Polish	500
59	Wymysorys	500
60	Wymysorys	500
61	Polish	500
62	Wymysorys	500
63	Polish	500
64	Wymysorys	500
65	Polish	500
66	Wymysorys	400
67	Wymysorys	400
68	Polish	400
69	Polish	400

Loop	Lang. Cue	Cue speed
70	Wymysorys	400
71	Wymysorys	300
72	Polish	300
73	Wymysorys	300
74	Wymysorys	300
75	Polish	300
76	Wymysorys	300
77	Wymysorys	300
78	Polish	300
79	Wymysorys	300
80	Wymysorys	300
81	Polish	300
82	Wymysorys	300
83	Wymysorys	400
84	Polish	400
85	Polish	400
86	Wymysorys	400
87	Wymysorys	400
88	Polish	500
89	Wymysorys	500
90	Polish	500
91	Wymysorys	500
92	Polish	500
93	Wymysorys	500
94	Wymysorys	500
95	Polish	500
96	Polish	500
97	Polish	500
98	Wymysorys	500
99	Wymysorys	500
100	Wymysorys	500

Appendix 2: Selected images from BOSS

accordion01
ant
apple07
barn
barnowl
bat
beaver
beermug02
birdnest
blackberry
bowrake
broom02
bucket01b
bun
button01
cabbage
canadiangoose
candle08b
castle
chair
cherry01
chimpanzee
chipmunk
cleaver02
clock03b
clothespin04
cloud
cockroach02
comb04
cookingpot
cow
crown
cucumber
doorlock
doublebass
duck01
feather03a
fence01
flipflop01c
fork03b
fryingpan03
grape01
grasshopper
greenonion
hammer01
hen
highheelshoe02a
hikingshoe01
honeybee
horse
huntingknife
key02a
ladle02b
ladybug03
lamp05
lettuce
lightbulb02a
lip
mirror02
monarchbutterfly
mug03
necklace
nose
onion
paintbrush01
pants
parkbench03
pencil02a
pig
pigeon
plate02b
potatomasher03b
raccoon
radiator
raspberry01
rawchicken
redfox
rockbass
rooster
rowboat
saw02b
screw01
sheep
shovel01
slipper02a
spoon01
stairs
strawbasket05
stringbean02a
toothbrush02a
tshirt
tulip02
wagonwheel
walnut01b
watch04
wheat
wheelbarrow01
workglove03a
worldmap
zipper

Appendix 3: distance matrix

Speaker	ms9907	mz0205	an0005	kz9902	bc8612	jm9303	tk9307
ms9907		0.6732634	0.1521456	0.9110996	0.2923490	0.6170418	0.0000000
mz0205	0.0197322		0.9893450	0.9952247	0.9998929	0.9999577	0.0002972
an0005	0.9953788	0.0028116		0.7002391	0.9986318	0.8622604	0.0011766
kz9902	0.7965697	0.4724633	0.4020251		0.9090586	0.9988316	0.0000001
bc8612	0.7443222	0.0000536	0.9827165	0.0549010		0.9855235	0.0000137
jm9303	0.0000000	0.0000173	0.0000000	0.0000000	0.0000000		0.0000000
tk9307	0.9503336	0.2365912	0.6555534	0.9996492	0.1467803	0.0000000	

Table 4: A speaker to speaker distance matrix of P values, comparing differential latencies of reaction times in Wymysorys and Polish. Wymysorys is represented by the upper triangle, and Polish by the lower.