Running head: PERCEIVED PHONETIC DIFFICULTY OF ENGLISH WORDS

Rule-based and Empirical Rating of Perceived Phonetic Difficulty of English Words According to Polish Learners: Does Frequency Matter?

Włodzimierz Sobkowiak

School of English, Adam Mickiewicz University

Abstract

208 Polish students of English philology filled in a questionnaire concerning the perceived phonetic difficulty of twenty English words stratified on two dimensions: (a) a-priori rule-based assessment of phonetic difficulty and (b) word frequency rank. A two-way ANOVA confirmed the significance of both main effects and their synergetic interaction, i.e. the perceived difficulty rating was affected by both the word's rule-based difficulty index and its frequency independently, as well as by their product. Conclusions are drawn for the theory and practice of EFL teaching in general, and for the design of EFL dictionaries with phonetic access in particular.

Rule-based and Empirical Rating of Perceived Phonetic Difficulty of English Words According to Polish Learners: Does Frequency Matter?

Foreign language learners approach their task with a large set of preconceived ideas: biases, stereotypes, attitudes, hang-ups. Some of these relate directly to the language which they are studying. A subset of these, in turn, concern the perceived difficulty of various structural and functional features of the language. Finally, some of this subset are subjective impressions of the phonetic difficulty (PD) of the foreign tongue under practice. These extend over both relevant skills (speaking and listening[1]) as well as over all areas of phonetics (from segments to intonation). Some representative examples of such PD judgements referring to English might be: “The first sound of *this* is the hardest one in English to pronounce,” “The English tend to swallow word endings, which makes it hard to understand them,” “I never know where a compound word should be stressed!,” “*Southern* is confusing in pronunciation because it is not similar to *south*.” Needless to say, such assessments are extremely L1-sensitive, so that the rating of perceived PD will be a product of inherent physio-articulatory complexity of L2 (which may, admittedly, be rather difficult to pinpoint) as well as L1-specific problems in mastering L2 pronunciation.

Foreign language teaching methodology seldom takes such learners' beliefs and attitudes into account. There are, I think, a number of mutually reinforcing reasons why this should be so. First, in a predominantly communicative paradigm of teaching/learning there tends to be little emphasis on phonetic accuracy. Second, a learner-centred classroom, with students autonomously following their self-set goals, is not a place where pronouncing problems and judgements would be high on the list of priorities. Third, teachers are often undertrained in phonetics themselves, which makes them wary of dissecting pronunciation niceties of, as they believe, little practical import to their pupils. Fourth, with the strong version of Contrastive Analysis Hypothesis (CAH) by and large discredited, phonetic transfer

and interference from L1 to L2 speech used as sole explanantia of pronunciation errors have come to be regarded with suspicion (see Dechert & Raupach, 1989; Dulay & Burt, 1978; Eckman, 1977; Flege & Davidian, 1984; Gass & Selinker, 1992; Hammarberg, 1997; Odlin, 1989).

It is encouraging, however, that these attitudes are beginning to change, at least as far as English-as-a-foreign-language (EFL) pedagogy is concerned. As noticed by Dalton & Seidlhofer (1994, p. ix), for example, in their book on *Pronunciation*, "Over recent years there has been renewed interest in the teaching of pronunciation which has resulted in a bewildering variety of new teaching materials being published." The 'new teaching materials' include pronunciation primers (which the two authors had in mind), but also vocabulary lists, dictionaries and multimedia software of all kinds (including internet-based).

According to some vocabulary learning specialists also lexical syllabi should be graded on the basis of the inherent difficulty of words selected for inclusion, including pronouncing difficulty. Nation (1990, p. 36), for example, suggests that at the beginning of the course teachers should introduce words phonetically 'easy' to learners of the given L1, with gradually more difficult words phased in later, as "Learning will be made easier because part of the learning burden—pronunciation—has been reduced." That it is in fact a burden in vocabulary learning was experimentally confirmed by Ellis and Beaton, who confirmed that "In learning the foreign vocabulary for native words, the pronounceableness of the foreign word has a strong determining effect (0.37) depending on the degree to which it conforms to the phonotactic patterns of the native language" (1993, p. 93). Papagno & Vallar (1995), Papagno, Valentine & Baddeley (1991), and Service (1992) associated this fact with the important role of short-term phonological memory in FL vocabulary learning.

Similar difficulty issues in a psycholinguistic context were raised as early as Lado (1955), to be followed by Higa (1965), Rodgers (1969), Richards (1974), Eckman (1981),

Ellis & Beaton (1993), Laufer (1990) and (1997), and de Groot & Keijzer (2000), among others. In this connection the availability of dependable English wordlists and dictionaries with PD rating for the given native tongue is dramatically underscored. Subjective ratings of such difficulty collected from EFL learners should be an indispensible basis for their compilation.

Phonetic Difficulty Index

The project which I report on below originated from these premises. A few years ago I was involved in the design of a radically innovative electronic dictionary of English for foreign learners. Part of the innovation was in the wide and flexible phonetic access to the contents of the dictionary, hence PAD, Phonetic-Access Dictionary. Briefly, the user (teacher, learner, materials developer, syllabus designer, test maker) could access practically all phonetic information inherent in the entry: (broad) transcription, syllable number and structure, stress pattern, segmental length, dialectal features and others. All these keys could function both representationally and indexically, i.e. as retrieval keys, singly and in (Boolean) combination. For example, a search over <a>-initial entries for iambic bisyllables containing a /-bs-/ cluster (hard for Poles on account of unassimilated voicing) would yield: *abscond, absorb, abstain, abstruse* and *absurd*.

One way in which PAD addressed the phonetic needs and problems of Polish learners of English was the consideration of common difficulties and errors which they face in acquiring the English phonological system. I decided to conflate these into a global measure of phonetic difficulty, an index (henceforth PDI), attached to each word. In the context of PAD the word was naturally the only feasible entity capable of carrying such an index.

The difficulty index could be used in a number of ways in the actual word searches and queries. First, it could caution the user as to the high PD of the word currently displayed. This piece of information could then be used in a variety of ways (practice, further lookup, exercises). Second, PDI could be used in direct interactive queries of the type: “Which words of this or that semantic/morphological category are particularly difficult phonetically?,” or “Give me the phonetically difficult words of the first 1000 (spoken) frequency rank.” Third, because the index in its full version would contain information about the specific type of the PD involved, it would allow the user to investigate it directly through listing words with this same difficulty, say a troublesome grapho-phonemic correspondence case. Fourth, the index could be used in the semi-automatic construction[2] of pronunciation exercises, which could thus be made sensitive to the inherent PD of the lexical items drawn from the database.

The index combined (a) the most salient grapho-phonemic difficulties such learners are known to have reading English, i.e. mostly spelling pronunciation; see e.g. Ellis & Beaton, (1993, p. 568-9); (b) some of the most common phonemic and phonotactic L1-based problems known from the literature and my own teaching experience; see e.g. Jassem (1973), Reszkiewicz (1981), Bałutowa (1999), Sobkowiak (2001); finally (c) some of the notorious developmental L2-interference pronunciation errors observed in all learners of English regardless of their L1 background; see Zobl (1980), Flege & Davidian (1984) or Hancin-Bhatt & Bhatt (1997). For details of its rationale, the process of its creation, including the choice of difficulties, the assignment of weights and the algorithm used, see Sobkowiak (1999), chapter three.

The range of points appearing in PDI was between zero and ten. To take one example showing how the 'rules' used in the PDI algorithm apply to one English word: *absorb* scored 4 points on the scale of difficulty (a medium score) for: (a) containing a schwa, (b) ending in a voiced obstruent, (c) having a lenis obstruent before /s/ and (d) differing significantly between

British and (General) American pronunciation (/r/-full spelling pronunciation for Polish speakers of RP).

In Sobkowiak (1999) I wrote:

> How seriously should one take these phonetic difficulty point assignments? It is certainly not the case that they reflect some absolute difficulty values which could be discovered theoretically or empirically from the analysis of Polglish[3] pronunciation. Naturally, L2 phonetic difficulty is just as dynamically developmental as the interlanguage itself. To be reliable, the difficulty index would have to be derived from the careful inspection of errors and the perceived problems of learners at different levels of proficiency. As a matter of fact, the present version of the index has so far not been tested on a group of Polish EFL learners. It is easy enough to imagine (but much more difficult to carry out!) a questionnaire or test which would target some of the PAD words from various bands of difficulty. The empirical results could then be matched against the scores derived from theory, intuition and personal experience. (Sobkowiak, 1999, p. 218)

In the remainder of this paper I report on a questionnaire study which I conducted to test just how well my intuitive assessment of the PD of some English words, ultimately supported by the Contrastive Analysis Hypothesis, fits the actual judgement of Polish EFL learners. The effect of word frequency will also be considered.

The Experiment

As the PAD difficulty index was developing, I was wondering about the possible effect of word frequency on the evaluation of PD. Specifically, I hesitated whether or not to include a condition in the PDI assigment algorithm adding points for low frequency,

regardless of other, strictly phonetic, criteria. It seemed intuitively plausible that the learner's familiarity with a given lexical item might affect their subjective assessment of the word's pronouncing difficulty. The learner would feel that s/he has not had much opportunity to practise the word, both receptively and productively, so s/he might feel uncertain as to the details of its pronunciation (as well as morphology, grammar, stylistic conditioning or pragmatics), and consequently might rate it as phonetically harder. There is also some less impressionistic evidence that phonetic and statistical markedness correlate, specifically that the incidence of phonetically marked segments (peripherals) increases with lexical frequency rank (Sobkowiak 2000b; see also Fenk-Oczlon, 2001, for a summary of relevant bibliography).

While there are reasons to believe that word familiarity proper would be a better measure of the intensity of contact that learners have with particular lexical items in the foreign language (Kacinik, Shears & Chiarello, 2000; Kreuz, 1987), I had to take its more easily available and quantifiable correlate, i.e. word frequency.[4] Word familiarity rating requires extensive experimentation on human subjects in rigorously controlled conditions, which explains why only a small proportion of English vocabulary has been so rated (6% in the MRC Psycholinguistic Database; see Coltheart, 1981; Wilson, 1988; and http://venus.cis.rl.ac.uk/; see also Gilhooly & Logie, 1980). Extracting appropriate bisyllables with required familiarity and PDI was not possible, let alone obtaining familiarity figures for all keywords of the planned PAD.

Thus, at the time when the PD index was tested against learners' judgements I finally decided that frequency would be one of the controlled variables. This decision affected the design of the experiment, of course, as two-way ANOVA would be unavoidable, with the main effects being the word's PDI and frequency.[5]

Subjects and Procedure

The questionnaire (see Appendix for a facsimile) was distributed in mid-February 2000 to the students of the first two years of English philology (both British and American) in the School of English, Adam Mickiewicz University, Poznań (roughly in the top hundred band on the paper-based TOEFL). It was applied indirectly, with the help of pronunciation teachers, to groups of students having their regular EFL practical phonetics classes. In the cover letter attached to a batch of questionnaire forms the teacher was asked to "run the attached questionnaire at your earliest convenience in the phonetics class you teach. It should not take more than a few minutes. The instructions to the questionnaire are simple so you should not be required to answer any questions from your students."

Students were asked to judge the words' PD to beginning Polish learners of English, i.e. not to themselves. This trick was used because in a small pilot study run on other School of English students it turned out that relativizing the question to me consistently produced a floor effect, i.e. all words turned out to be (suspiciously) easy phonetically. Asking about 'beginning learners' does not threaten the respondent's face while at the same time yielding a personalized answer, nevertheless. Practically, I assume, the question most respondents were answering was: "how difficult was this word to me when I was a beginner?," which is adequate for my purposes.

208 returns were collected, which is about 90% of all 1/2 year students. These were converted into a computer-readable form for further processing; a spreadsheet of 20 (words) x 208 (respondents) = 4160 cells originated, each containing an integer between 1 (phonetically easy word) and 4 (difficult). No missing data points were noticed. The occasional frivolous returns were not filtered out (e.g. there were ten returns with all twenty words graded 1 and one return with maximum grades throughout).

Data

To select the twenty words to be phonetically graded by the respondents two word lists were correlated. One was derived from PAD, as described above in section 2, and contained bisyllables stratified into four PDI levels, as measured by the index: 0, 2, 4 and 6, from phonetically easiest to hardest. Bisyllables only were used to reduce the amount of uncontrolled variation in potentially phonetically relevant properties of the stimulus words, such as segmental (phonemic and graphemic) length, stress pattern or consonant cluster incidence. It would not be feasible to stringently control for all of these variables on top of the two tested in the experiment, PDI and frequency.

Only four levels of relatively low PDI were taken into account, with 2-point intervals because: (a) few phonetically difficult bisyllables appeared in the list, (b) it was felt that 1-point intervals would not give satisfactory discrimination of PDI, (c) it was straightforward to match a four-point PDI scale with the four-point scale imposed on the respondents.

The other list was the lemmatized frequency list of 6318 words (those with frequencies of 800 upwards) derived by Adam Kilgarriff from the 100-million-word-long British National Corpus (for further information on BNC, see http://info.ox.ac.uk/bnc/; to download Kilgarriff's list go to ftp://ftp.itri.bton.ac.uk/pub/bnc/; see also Kilgarriff, 1997). Five frequency rank bands were selected to generate twenty stimulus words altogether; this number was believed to be about right in a small scale questionnaire to be run as part of ordinary classes. The five rank bands narrowly clustered in equal intervals of one thousand around 300, 1300, 2300, 3300 and 4300.[6] These levels were selected because: (a) the first couple of hundred rank words on any frequency list are practically all function words and content-word monosyllables, (b) one thousand rank interval provides for good discrimination,

(c) it is increasingly more difficult to find properly PDI-stratified bisyllabic words at higher frequency ranks.

In the questionnaire presented to the respondents all words were randomized. From the provided intructions the subjects could guess that the words were in fact PDI-stratified (see Appendix), but no mention of the frequency variable was made. With all these experimental controls there is a fair chance that the results are a representative reflection of Polish EFL learners' subjective PD judgements (henceforth PDR, for PD Rating) relative to the rule-based PAD PD indexing and word frequency.

Results and discussion

Averaged results of the questionnaire are tabulated in Table 1, the PDI across the top, and the frequency rank down the left-hand side. Each of the twenty cells contains the stimulus word and the learner-assigned PDR value averaged over 208 returns, followed by standard deviation. Thus, for example, the phonetically easiest of the twenty words, as judged by the 208 respondents, is *taxi*, with the mean PDR of 1.2 (on a scale of 1–4) and sd = 0.46; the hardest is *southern* — 3.2, with sd = 0.90.

Global means are shown at the bottom and in the rightmost column of the table. It will be seen that students' ratings of PD shows a good correlation with my assessment expressed as PDI: both values rise consistently from left to right. Similarly, if less expectedly and consistently, students' ratings correlate with the word's frequency rank: the higher the rank the higher PDR, with the trough at rank 3300 (*taxi* and *server* appeared to be much easier phonetically than the other two words of this rank). This apparent double conditioning of PDR by both PDI and rank is graphically represented by the shading of the cells, increased in 0.20 PDR increments, globally getting diagonally darker (harder) from top-left to bottom-right.

Table 1 here

I used the Pearson product-moment test to measure correlation of my rule-based PD index with the empirical students' mean PDR over the entire list of twenty words. As it turned out, the respondents' difficulty ratings correlate very highly with my PDI (r = .684, r^2 = .468 p <.0005 for df = 19, one-tailed). This is best seen in Figure 1, where a scatterplot of the two is presented with the best-fitting regression line and equation.

Figure 1 here

Thus, we are now a bit closer to answering the original question posed in Sobkowiak (1999, p. 218): “How seriously should one take these phonetic difficulty point assignments?” (made on the basis of relevant literature as well as researcher's experience and intuition). The answer appears to be: quite seriously, because they have a good empirical basis and correlate highly with empirically obtained learners' judgements. In other words, a suitably calibrated computer algorithm assigning L1-sensitive phonetic difficulty indexes by rule to English words is a reliable tool to enhance the content of EFL dictionaries, word lists and other teaching materials and resources. This is notwithstanding the obvious need for continuous improvement of the match between the learners' PD rating and the PD index.

So, finally, should word frequency be taken into account in this latter task? A tentative 'yes' appeared from the cursory analysis of Table 1: less frequent words appeared to be rated as more phonetically difficult, other things being equal (to the extent that they can).

Correlation testing, however, was disappointing, with $r = .16$ ($r^2 = .026$), which is not significant at $df = 19$.

As the frequency-related results so far were equivocal, I decided to subject the data to the analysis of variance. A two-way repeated measure factorial experimental design was used with one dependent variable (PDR) and two experimental factors, PDI and rank, with four and five levels, respectively. The results appear in Table 2.

Table 2 here

Table 2 shows how the two chosen factors (PDI and rank) affect the level of the dependent variable, i.e. PDR assigned by learners. Both main effects are significant. Ratings of perceived difficulty were higher for words with a high PD index ($F(3, 4152) = 249.01$, p«.01). Ratings of difficulty were likewise higher for words of higher frequency rank ($F(4, 4152) = 49.61$, p«.01). As the critical level of $F(3, \infty)$ at $p = .01$ equals 3.78 and $F(4, \infty)$ — 3.32, it is immediately obvious that both effects in this analysis are significant statistically.

Thus, results indicate a correlation between learners' PD ratings and my PDI, the conclusion reached earlier which is here confirmed by another test. Additionally, and perhaps more interestingly, ANOVA shows that learners, in their PD ratings, are also sensitive to word frequency: the higher the frequency (the lower the rank) of the word the easier phonetically it is graded. The slightly lower level of the F statistic for this effect is doubtless due to the slump around rank 3300 mentioned earlier. Finally, notice that the interaction between the two main effects is also significant, with $F(12, 4152) = 84.00$ (the critical level at $p = .01$ being $F(12, \infty) = 2.18$). This means that the rating of phonetic difficulty grows synergetically as a product of both main effects in direct proportion to their combined

influence, as approximately shown by shading in Table 1. Another graphic way in which this effect can be rather dramatically illustrated is by feeding the data into a three-dimensional diagram (Figure 2) shown below, where the shaded plane representing PDR rises diagonally towards the high-PDI/high-rank corner.

Figure 2 here

All these results have rather direct consequences for the design of the PAD phonetic difficulty index. First, it seems that it is a viable idea of a tool significantly enhancing the functions of learner dictionaries, both traditional and electronic. Second, while the correlation of the currently generated PD index with learners' ratings is high, it can of course be made even higher. It is now clear how empirical data can be used in this process. Third, word frequency is an important factor in the learners' assessment of the word's phonetic difficulty, at least in the context of the present experiment: English as a foreign language in Poland, the (rank-wise) first few thousand bisyllabic lemmas, the used sample of informants, etc. How it will contribute to the global PD index is another matter, but these are the technicalities of the algorithm which can be adjusted to fit empirical data as much as possible.

Conclusions

Word frequency does matter in the subjective judgement of the word's phonetic difficulty by foreign learners. The more common the word the less phonetically difficult it seems, even if its rule-derived PD index and a few other phonetically sensitive variables are kept constant. This in itself is not a revolutionary finding: educators, language teaching methodologists, psycholinguists, lexicographers, statisticians of language and corpus linguists

have long since discovered the crucial role played by frequency and markedness in natural languages. But the effect of lexical frequency on the perceived phonetic difficulty has not so far been considered. As a matter of fact, it has been of only marginal interest in constructing syllabi and word lists in the lexical approach to foreign language teaching, one which has taken word frequencies more seriously than others.

This should now change. Both pronunciation and frequency are quite simply too important aspects of vocabulary structure and function to be left outside of the realm of interest of language teachers, syllabus creators, material designers, lexicographers, translators, NLP specialists and a host of applied linguists of many different denominations. If they are ignored, the resulting insights and resources will exhibit clear faults, which might ultimately harm learners. This I demonstrated on the example of phonetic keywords (Sobkowiak, 2000a) and defining vocabularies (Sobkowiak & Kuczyński, 2002) in EFL dictionaries.

In this paper another tiny niche of the huge issue was inspected, the interaction of word frequency and its perceived phonetic difficulty by foreign learners. No mention has been made of other linguistic units where the two aspects might affect each other: the phoneme, the morpheme, the phrase, the intonation contour. Also, a rather small and specific group of respondents participated, and it is probable that results obtained from other populations (secondary school learners, teachers, natives) would have been different. The field is now open for analysis. It is to be hoped that the benefits will be reaped by the learner.

References

Bałutowa, S. (1999). *Wymowa angielska dla wszystkich*. Warszawa: Philip Wilson.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A,* 497-505.

Dalton, C., & Seidlhofer, B. (1994). *Pronunciation*. Oxford: Oxford University Press.

De Groot, A.M.B., & Keijzer, R. (2000). What is hard to learn is easy to forget: the roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning, 50.1,* 1-56.

Dechert, H.W., & Raupach, M. (Eds). (1989). *Transfer in language production*. Norwood, NJ: Ablex Publishing.

Dulay, H., & Burt, M. (1978). Some remarks on creativity in language acquisition. In W. C. Ritchie (Ed.), *Second language acquisition research: issues and implications* (pp. 65-89). New York: Academic Press.

Eckman, F.R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27,* 315:30.

Eckman, F.R. (1981). On predicting phonological difficulty in second language acquisition. *Studies in Second Language Acquisition, 4,* 18-30.

Ellis, N.C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43,* 559-617.

Fenk-Oczlon, G. (2001). Familiarity, information flow, and linguistic form. In J. Bybee, & P. Hooper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 431-48). Amsterdam: Benjamins.

Flege, J.E., & Davidian, R.D. (1984). Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics, 5,* 323-47.

Gass, S.M., & Selinker, L. (Eds). (1992). *Language transfer in language learning*. Rowley, MA: Newbury House.

Gilhooly, K.J., & Logie, R.H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behavior Research Methods and Instrumentation, 12,* 395-427.

Hammarberg, B. (1997). Conditions on transfer in phonology. In A. James, & J. Leather (Eds.), *Second language speech. Structure and process* (pp. 161-80). Berlin: Mouton de Gruyter.

Hancin-Bhatt, B., & Bhatt, R.M. (1997). Optimal L2 syllables: interactions of transfer and developmental effects. *Studies in Second Language Acquisition, 19,* 331-378.

Higa, M. (1965). The psycholinguistic concept of 'difficulty' and the teaching of foreign language vocabulary. *Language Learning, 15,* 167-79.

Jassem, W. (1973). *Podręcznik wymowy angielskiej*. Warszawa: Państwowe Wydawnictwo Naukowe.

Kacinik, N., Shears, C., & Chiarello, C. (2000, August). Familiarity for nouns and verbs: not the same as, and better than, frequency. Paper presented at the 22nd Annual Meeting of the Cognitive Science Society, University of Pennsylvania. Retrieved November 16, 2001, from http://www.cis.upenn.edu/~ircs/cogsci2000/PRCDNGS/SPRCDNGS/abstrcts/kacsh_ch.pdf

Kilgarriff, A. (1997). Putting frequencies in the dictionary. *International Journal of Lexicography, 10.2,* 135-55.

Kreuz, R. J. (1987). The subjective familiarity of English homophones. *Memory and Cognition, 15,* 154-168.

Lado, R. (1955). Patterns of difficulty in vocabulary. *Language Learning, 6,* 23-41.

Laufer, B. (1990). Ease and difficulty in vocabulary learning: some teaching implications. *Foreign Language Annals, 23.2,* 147-55.

Laufer, B. (1997). What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 140-55). Cambridge: Cambridge University Press.

Nation, P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

Odlin, T. (1989). *Language transfer: cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30,* 331-47.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology, 48A,* 98-107.

Reszkiewicz, A. (1981). *Correct your English pronunciation*. Warszawa: Państwowe Wydawnictwo Naukowe.

Richards, J.C. (1974). Word lists: problems and prospects. *RELC Journal, 5.2,* 69-84.

Rodgers, T. (1969). On measuring vocabulary difficulty: an analysis of item variables in learning Russian-English vocabulary pairs. *International Review of Applied Linguistics, 7,* 327-43.

Service, E. (1992). Phonology, working memory, and foreign language learning. *Quarterly Journal of Experimental Psychology, 45A,* 21-50.

Sobkowiak, W. (1999). *Pronunciation in EFL machine-readable dictionaries*. Poznań: Motivex.

Sobkowiak, W. (2000a). Phonetic keywords in learner's dictionaries. In U. Heid, S. Evert, E. Lehmann, & C. Rohrer (Eds.), *Euralex'2000 Proceedings* (pp. 237-46). Stuttgart: Institut für Maschinelle Sprachverarbeitung.

Sobkowiak, W. (2000b). When are peripheral plosives preferred?. *Studia Phonetica Posnaniensia, 6*,97-107.

Sobkowiak, W. (2001). *English phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.

Sobkowiak, W., & Kuczyński, M. (2002). The phonetics and ideology of defining vocabularies. Manuscript in preparation.

Suzuki, Y. et al. (2001). A new method to assess the intelligibility in daily speech communication. Towards the best fitting of hearing aids. *NTT News Release, 10th September 2001*. Retrieved December 10, 2001, from http://www.ntt.co.jp/news/news01e/0109/010910.html

Wilson, M. (1988). MRC psycholinguistic database: machine-usable dictionary, version 2.00. *Behavior Research Methods, Instrumentation and Computers, 20.1*,6-10.

Zobl, H. (1980). Developmental and transfer errors: their common bases and (possibly) differential effects on learning. *TESOL Quarterly, 14.4*, 469-83.

Appendix

Phonetic difficulty questionnaire

Dear Respondent!

This is a subjective phonetic difficulty questionnaire. It is part of my research into how Polish learners of English explicitly grade the subjectively perceived pronuncing[7] difficulty of certain English words. Apart from the inherent scientific interest, the research is driven by the hope of being able to improve English teaching resources and materials: textbooks, tests, dictionaries, etc.

The success of this *anonymous* questionnaire crucially depends on your frank and spontaneous evaluation of *pronouncing difficulty* of the twenty words below. Notice that I am *not* interested in your opinion about the word's semantic or grammatical difficulty. For each word you should ask yourself the same question: **how difficult to pronounce is this word *to beginning Polish learners of English*, on the scale from 1 (easy) to 4 (hard)?**

You will be informed of the global results of the questionnaire when they are ready.

Thank you for your time and effort!

author	**....**	**oblige**	**....**	**defect**	**....**	**southern**	**....**
tired	**....**	**survive**	**....**	**coloured**	**....**	**relax**	**....**
taxi	**....**	**mother**	**....**	**dissolve**	**....**	**appear**	**....**
debate	**....**	**almost**	**....**	**belief**	**....**	**youngster**	**....**
carry	**....**	**awkward**	**....**	**server**	**....**	**kingdom**	**....**

Author Note

Włodzimierz Sobkowiak, School of English, Adam Mickiewicz University, Poznań, Poland.

Correspondence concerning this article should be addressed to Włodzimierz Sobkowiak, School of English, Adam Mickiewicz University, ul. Niepodległości 4, 61-874 Poznań, Poland. E-mail: sobkow@amu.edu.pl.

I am grateful to the participants of the Phon&Phon meeting on March 30th 2000, for their feedback to my first presentation of the results of my project, and to dr Robert Lew for his generous help with the statistical aspects of the study.

Footnotes

[1] But also reading, if grapho-phonemic correspondences in the language are complex and/or erratic, as they are in English, for example.

[2] I.e. once the overall schema of the exercise is fixed by the author (teacher, developer, tester), words could be drawn automatically from a pool satisfying certain pre-set phonetic conditions, orthographic and morphological parameters, frequency/familiarity ranges, etc.

[3] I.e. the Polish-English interlanguage.

[4] See Benjamin K. Tsou et al.'s 1998 City University of Hong Kong project described on http://www.rcl.cityu. edu.hk/research/nl4_rh1.htm, wherefrom I took this quote: "Words with high frequency of occurrence are judged to be easier than those with lower frequency" (accessed 8 December 2001). Also Kreuz and Kacinik et al. admit that word frequency correlates very highly with reported familiarity, as high as r = 0.75 (Kreuz).

[5] A similar design was used by Suzuki et al. (2001) in their research on word intelligibility tests of hearing-impaired listeners. Lexical familiarity and phonetic structure were used as independent variables, the former as an index of word difficulty.

[6] The actual BNC frequencies of the twenty lemmas were as follows: almost 31588, appear 30595, author 6852, awkward 1489, belief 7509, carry 31258, coloured 1562, debate 7520, defect 1488, dissolve 1472, kingdom 3781, mother 27784, oblige 2200, relax 3768, server 2157, southern 4553, survive 7398, taxi 2184, tired 3762, youngster 2260.

[7] This typo appeared in the original form.

Tables

Table 1.

Mean/sd Learner PD Ratings (PDR) of Twenty Words, as a Function of Rank and PDI

	Phonetic Difficulty Index (PDI)								**mean**
≈ rank	**0**		**2**		**4**		**6**		**PDR**
300	carry	1.5/0.68	almost	1.8/0.88	appear	1.9/0.81	mother	2.1/0.96	1.8
1300	belief	1.5/0.68	debate	1.7/0.72	survive	2.1/0.83	author	2.9/0.96	2.1
2300	relax	1.5/0.65	kingdom	2.2/0.94	tired	1.9/0.89	southern	3.2/0.90	2.2
3300	taxi	1.2/0.46	oblige	2.5/0.99	server	1.8/0.79	youngster	2.1/0.88	1.9
4300	defect	1.7/0.73	dissolve	2.0/0.84	awkward	3.1/0.93	coloured	2.1/0.77	2.2
mean PDR		1.5		2.0		2.2		2.5	

Note. Shading of the cells is increased in 0.2 PDR increments.

Table 2.

Two-way ANOVA for Phonetic Difficulty Index (PDI) and Rank

effect	**df**	**MS**	**MS error**	**F-ratio**
PDI	3	186.25	0.75	$F_{3, 4152} = 249.01$**
Rank	4	25.34	0.51	$F_{4, 4152} = 49.61$**
PDI x Rank	12	41.28	0.49	$F_{12, 4152} = 84.00$**

***P* <.001

Figure Captions

Figure 1. Correlation of Phonetic Difficulty Index (PDI) and Phonetic Difficulty Rating (PDR).

Figure 2. Synergetic effect of PD index and frequency rank on PD rating.

Figure 1.

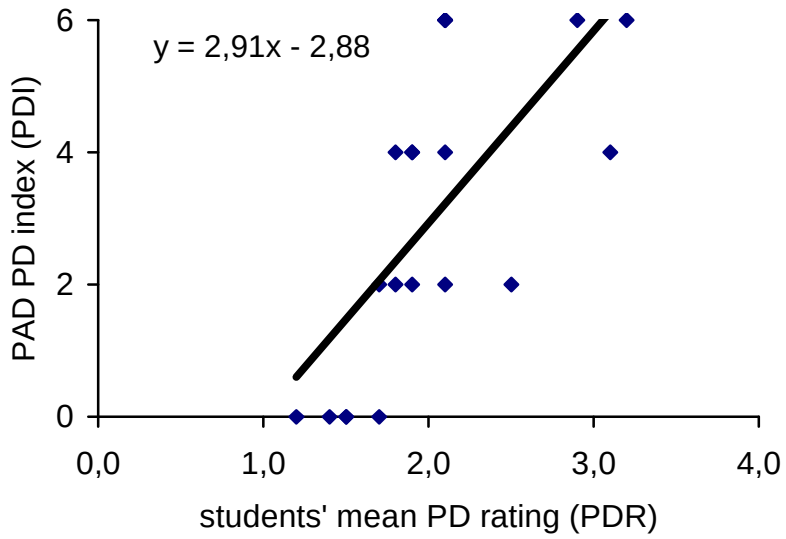
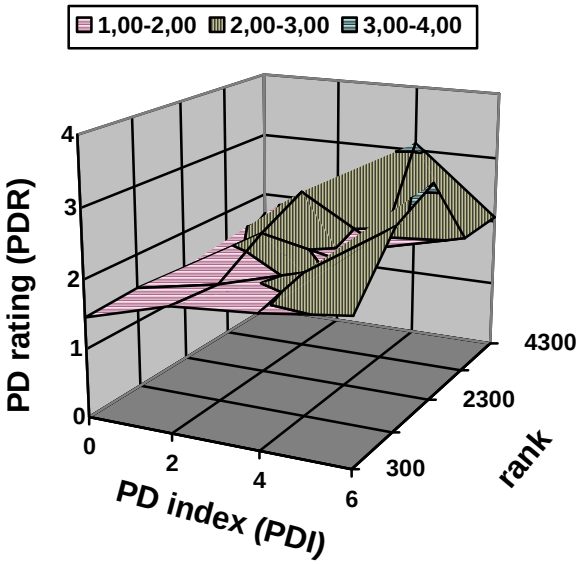

y = 2,91x - 2,88
PAD PD index (PDI)
students' mean PD rating (PDR)
0
2
4
6
0,0
1,0
2,0
3,0
4,0

Figure 2.


1,00-2,00
2,00-3,00
3,00-4,00
PD rating (PDR)
4
3
2
1
0
0
2
4
6
PD index (PDI)
4300
2300
300
rank

Notes

[1] But also reading, if grapho-phonemic correspondences in the language are complex and/or erratic, as they are in English, for example.

[2] I.e. once the overall schema of the exercise is fixed by the author (teacher, developer, tester), words could be drawn automatically from a pool satisfying certain pre-set phonetic conditions, orthographic and morphological parameters, frequency/familiarity ranges, etc.

[3] I.e. the Polish-English interlanguage.

[4] See Benjamin K. Tsou et al.'s 1998 City University of Hong Kong project described on http://www.rcl.cityu.edu.hk/research/nl4_rh1.htm, wherefrom I took this quote: “Words with high frequency of occurrence are judged to be easier than those with lower frequency" (accessed 8 December 2001). Also Kreuz and Kacinik et al. admit that word frequency correlates very highly with reported familiarity, as high as r=0.75 (Kreuz).

[5] A similar design was used by Suzuki et al. (2001) in their research on word intelligibility tests of hearing-impaired listeners. Lexical familiarity and phonetic structure were used as independent variables, the former as an index of word difficulty.

[6] The actual BNC frequencies of the twenty lemmas were as follows: almost 31588, appear 30595, author 6852, awkward 1489, belief 7509, carry 31258, coloured 1562, debate 7520, defect 1488, dissolve 1472, kingdom 3781, mother 27784, oblige 2200, relax 3768, server 2157, southern 4553, survive 7398, taxi 2184, tired 3762, youngster 2260.

[7] This typo appeared in the original form.