INVESTIGATING IMITATIVE ABILITY

Duncan Markham
Department of Linguistics and Phonetics, Lund University, Sweden

ABSTRACT

This paper introduces a newly initiated research project dealing with imitation, entitled Human Imitation: Perceptual and Productional Processes. The project is motivated by the poor understanding of phonetic behaviour and ability in second language acquisition, and voluntary adoption of non-native dialect, accent and speaker characteristics.

INTRODUCTION

At present there exists little conclusive research to explain why phonetic acquisitional capacity (or any other acquisitional capacity in linguistics) appears to deteriorate with age. It is generally assumed, and has been empirically demonstrated, that both perceptual discrimination of unfamiliar phonetic contrasts (eg [1]) and productional ability (accentedness in L2) (eg [2]) do deteriorate with age, but there have also been investigations indicating no age effect (cf [3, 4]). Other investigations have demonstrated that general multilingual experience does not facilitate imitative/acquisitional ability [5], contrary to another common belief. Much of the published research, including some mentioned above, draws powerful conclusions from comparisons with other research in phonological acquisition, despite often grossly different research methods and controls. This research project attempts to take into account as many of the potential influencing factors as possible, and specifically addresses the question of individual ability — an issue ignored almost entirely in phonological performance studies.

PILOT INVESTIGATION

The results of a pilot investigation of imitative ability will be presented here. The investigation served as a test of the elicitation methodology and experimental design to be applied in the main investigation. Our primary interest in the pilot investigation was the experimental structure and the reactions of the informants, rather than a detailed investigation of their performance. This was necessary, as the design places considerable demand on the speakers' ability and concentration, and involves a large amount of preparation and training. In the following section, the investigation and its results will be described. The structure of the main investigation will then be described.

A text, approximately semantically equivalent, was prepared for a number of languages. Each text was read by at least one native speaker at a slow speed and at a faster casual/colloquial speed. A word list was then read once by each speaker. The word list contained all words from the text, each word appearing twice in the list. Two minutes of spontaneous speech were also recorded for each speaker (in most cases, the speaker described his house). Native speakers were selected on the basis of age (20-45) and modal voice, as far as practicable, as the imitators were to be of similar age (20-35), and in order to maximise the number of phrases which could be extracted as stimuli (see below).

The recordings were made on DAT at a sampling rate of 44.1kHz. They were then resampled using the SoundScope hardware/software package on a Macintosh computer at 22.2kHz. The texts were examined for short phrases or subphrases suitable for use in the experiment. The phrases had to be syntactically whole, in that constituents couldn't be divided. For example, in the phrase *the linguist parses the sentence*, the subphrases *the linguist, the linguist parses, parses the sentence, the sentence* would be acceptible, whilst **linguist parses, *linguist parses the, *parses the, etc.* would not. Furthermore, the phrases had to exhibit fairly modal voicing characteristics. Syntactically conditioned creak was acceptible, but speakers who exhibited strong creak as part of their reading register in general, or in their lower range for F0, were not used. The phrases were excerpted from the texts at both speeds, as were the relevant words

*short tone (450Hz, 125ms) // three instances of a fast phrase eg, *the linguist parsed the sentence*;
*short tone // three instances of the slow phrase;
*two-tone signal (450Hz, 125ms; 600Hz, 125ms) // two instances of the first grammatical word in the phrase (from the word list) eg, *the ... the*;
*two-tone signal // two instances of the second grammatical word in the phrase;
...
*two-tone signal // two slow instances of the first subconstituent eg, *the linguist ... the linguist* ;
*two-tone signal // two fast instances of the first subconstituent;
...
*two-tone signal // two slow instances of the phrase;
*two-tone signal // two fast instances of the phrase.

Figure 1. Stimulus presentation.

from the word lists.

A battery of stimuli were prepared for each language. For each phrase (two per language), the informants would hear complete phrases, and then the individual words or subconstituents, as shown in Figure 1. The stimuli were then recorded to DAT again, preceded by the fast text, the slow text, and (for some languages) the spontaneous speech.

Two informants (one male, one female) were chosen for the pilot investigation, on the basis of their above average performance in the pronunciation of some foreign languages. It was, however, necessary to choose informants of lesser apparent imitative ability than those participating in the main investigation, in order to maximise the number of good imitators remaining for the latter investigation (informants could not participate in both the pilot and main experiments). The informants were native speakers of Swedish, 26 and 29 years old respectively. The male spoke Southern Swedish (*skånska*), whilst the female spoke with a mixed regional accent of Halland, on Sweden's west coast (*halländska*) and *skånska*.

Three weeks before the experiment, the informants received an audio cassette containing the read texts and spontaneous speech for a selection of the languages. They were required to listen to the cassette twice on ten separate occasions (2 x 10=20 listenings) during the three weeks. The informants then came to the Department for the experiment. Recordings were made in the Department's recording studio.

Prior to the imitation task, the informants were fitted with a lightweight (30 gm) headset microphone, and asked to read a small part of the text used above (and with which they were, by now, familiar) in a dialect other than their own. They were also asked to read text excerpts, in as native-like a manner as possible, for languages in which they had adequate competence. For all dialects or languages they were also asked to introduce themselves as fictional persons, according to details they were given on paper (eg, Johnny, 29, London, student, photography, New York). These tasks were used to gain an impression of each informant's a) delayed imitative ability (strength of existing dialect models), b) base pronunciation performance in L2s, and c) performance in a non-reading (more freely structured) task.

For Swedish, the texts and introductions were done with regional pronunciation for Stockholm, and Finland-Swedish, with which all Swedes can be said to have some familiarity. For English, 'British English' and 'American English' were specified. For German and French no specific dialects were requested.

The informant was then told of the structure of the imitation task, and how the stimuli would be presented. This was done using visual information — pictures of the approximate sequence of signal-tones, texts, phrases, etc. The stimuli were then presented over high-quality headphones. The informant could not see the author, but communication was possible via a loudspeaker in the recording studio. After each two-tone signal, the informant heard two instances of a stimulus and then attempted to imitate it. Adequate time was available for each imitation.

After completing this task (total duration: 32 mins), informants were permitted to leave the studio, and were provided with some refreshments. After approximately 10 mins, the imitation task was repeated.

Canadian English (female, 24, Toronto) — TRAINING LG
have to work nights: have, to, work, nights
Japanese (female, 43, Osaka) — TRAINING LG
Sutefan wa pailotto de (Steven is a pilot): Sutefan wa, pailotto de
futari wa Nihon Kookuu de hataraite imasu (they both work for Japan Airlines): futari wa, Nihon Kookuu de, hataraite imasu, futari was Nihon Kookuu de, Nihon Kookuu de hataraite imasu (whole phrase not imitated, due to length)
Finland-Swedish (female, ca 40, Helsinki)
men omväxlande arbete (but varied job): men, omväxlande, arbete
till många storstäder runt om i världen (to many cities all over the world): till, många, storstäder, till många storstäder, runt, världen, runt om i världen
French (male, 45, Grenoble, no strong regional accent) — TRAINING LG
au deuxième étage (on the second floor): au, deuxième, étage
c'est pareil pour Maria (the same is the case for Maria): c'est, pareil, pour, Maria, c'est pareil, pour Maria
Mandarin (male, ca 50, Beijing)
dōu yǒu qìchē (each has a car): dōu, yǒu, qìchē
sānge fángjiān (3 rooms): sān, ge, fángjiān

Figure 2. Language and stimuli used.

RESULTS

Text Reading and Self-Introduction

The male informant, JE, showed interesting behaviour for the two Swedish dialects and two English dialects elicited. The text and introduction had been designed to highlight allophonic differences between speakers' native dialect and the imitated dialect. For JE, the clearest allophonic changes required for Stockholm Swedish were [ɧf] → [ʂ/ɧx], [ʀ] → [r/ʐ], [ᵊi:] → [i:]. and a change in pitch accent realisations. The first allophonic change, and the pitch accent change, were not attempted, which was surprising as they are considered to be highly salient dialect cues. For Finland-Swedish, on the other hand, JE made attempts to produce dialect specific [r ʃ y: ʉ:], typical breathiness and pitch accent, but not [tɕ] for [ɕ]. His productions of British English (attempted RP) sounded non-native, primarily because of the realisation of /ai/ and /a:/ with [ɒ] instead of a more front a-vowel. American English was more acceptable, though problems with the diphthong just mentioned were also observed. JE's normal English production tends towards RP, but residence in an English speaking environment was limited to the USA.

The female speaker, PO, showed much less resistance to pitch accent change required for three Swedish dialects (*skånska* was added), but she didn't modify *intonational* behaviour. Her segmental production showed clear and fairly good attempts at dialect-allophony for all dialects. PO's productions of both varieties of English were at times characterised by typically Swedish dental stops, but most other segmental and prosodic requirements were well met. Both informants had problems with reduction in both American and British English. The text included the sequence *so that they can*, and proved exceedingly difficult, as they tried to produce each syllable and fricative, rather than reducing the second and fourth syllables, which would normally result in something resembling [sᵊʊˈðɛɪkʰn̩]. JE's productions were always less native-like and sounded more exaggerated than PO's for all dialects.

Imitation

The two informants behaved at times quite differently in this task. JE behaved very poorly with regard to intonation reproduction, even for single word stimuli. There was slight improvement on the second trial. PO was considerably better, though still showed many errors. Her weakest area, however, was vowel quality, contrasting with fairly good performance for consonants. JE was far more variable, though generally not as good as PO, as he had difficulties with segmental production, utterance length (tempo), and rhythmical characteristics in French and Japanese (especially gemination in the latter — unexpected, as Swedish makes use of gemination contrasts).

Of considerable interest was the way in which intonation in phrases was usually reproduced. Generally, the onset of imitation was good, but the highly salient final tonal movements were not. This is curious, considering the perceptual salience of final tonal movements for expressing both syntactic and pragmatic functions, such as questions, scepticism, sadness, etc. Similar behaviour was observed for imitation of prosody only (rather than the segmental characteristics of an utterance) in a previous study [6].

MAIN INVESTIGATION

The main investigation takes essentially the same form as the pilot investigation. There are approximately eight informants (native speakers of Swedish), though it is hoped that further informants may be located. All informants show near-native ability in the pronunciation of at least one second language (L2). It is relatively difficult to find and recruit speakers who are this good and are prepared to be made aware of their imitative deficits, as language, and especially phonetic, ability is a sensitive part of the human ego.

Speakers will be screened prior to the training phase, so as to gain an impression of their present abilites. The screening involves both the initial reading and introduction task described above, and some imitative tasks. Shortly thereafter, half of the informants will receive an audio tape containing training material. This material will be different for each informant, as it will not include L1 or L2 dialects in which a given informant is already competent. The remaining informants receive no training.

The imitation task will take a similar form to the pilot investigation. The number of languages/dialects will, however, be increased to six. Additional imitations of the same stimuli will be made after the first attempt, rather than presenting the entire battery again. This means that rapid improvement for imitations of the same stimulus might be observable, as informants can monitor and correct their productions with little delay. This should also reveal the stimulus characteristics which remain unidentified or unproduced by a given informant, more readily than the delayed 'second chance' given in the pilot experiment. After the phrase tasks for a given language, informants will also be presented with a number of extra-contextual word stimuli, and phonological contrast-pairs (eg chair~share), for allophones or phonetic contrasts which are known to be problematic for speakers of Swedish (eg, alveolars, voicing in fricatives, syllabic consonants).

The elicited material will be edited and test tapes will be prepared for the next phase of the investigation. Trained phoneticians will assess the closeness of the imitations, and degree of non-native accent, auditorily (including discrimination tests — accented~unaccented) and attempt to describe deviations. The material will then be analysed instrumentally to assess acoustic deviation and its relationship to perceived deviation.

REFERENCES

[1] J. F. Werker and R. C. Tees (1984), Cross-language speech perception: Evidence for perceptual reorganization during the first year of life, *Infant Behavior and Development*, vol. 7, pp. 49-63.

[2] S. Tahta, M. Wood and K. Loewenthal (1981), Foreign accents: Factors relating to transfer of accent from the first language to a second language, *Language & Speech*, vol. 24, pp. 265-272.

[3] J. E. Flege (1987), The production of 'new' and 'similar' phones in a foreign language: Evidence from the effect of equivalence classification, *Journal of Phonetics*, vol. 15, pp. 47-65.

[4] C. E. Snow and M. Hoefnagel-Höhle (1977), Age differences in the pronunciation of foreign sounds, *Language & Speech*, vol. 20, pp. 357-365.

[5] J. F. Werker (1986), The effect of multilingualism on phonetic perceptual flexibility, *Applied Psycholinguistics*, vol. 7, pp. 141-156.

[6] D. J. Markham (1994), Prosodic imitation: Productional results. In ICSLP'94, vol. 3, pp. 1187-1190, Yokohama: Acoustic Society of Japan.