2 Transcription

De la Condamine says of a small nation living on the Amazon River: "Some of their words could not be written, not even most imperfectly. One would need at least nine or ten syllables where in their pronunciation they appear to utter hardly three."

Johann Herder, *On the Origin of Language*

1. Introduction

For most beginning linguists, transcription probably causes more headaches than any other task. This is primarily due to the well-known fact that it is not possible to identify word boundaries in an unknown language, and it is often impossible even to identify individual sounds. As many ancient writing systems reflected in their orthography, connected speech generally does not contain many word boundaries; the only thing that enables speakers of a language to identify word boundaries is their knowledge of the words of the language. Unfortunately, beginning field linguists generally do not know the words of the language they are studying, and therefore face an uphill struggle when trying to break down a stream of connected speech into individual words. We believe that it is a good idea for the professor in a Field Methods course to have the students try to carry out this task, just so they get an idea of how speech processing actually works (and doesn't work). Once this demonstration is completed, though, the students (and, even more so, linguists conducting field work on their own) are still faced with the challenge of identifying word boundaries that do not even exist in the speech signal.

There are several ways to make this task easier. For at least the first few sessions, collect words in isolation (for specific words to collect, see chapter 3; remember that what is one word in English may correspond to more than one word in your informants' language, and vice versa). This puts off the problem of identifying word boundaries until you have a better grip on the vocabulary of the language. While you are building up a basic vocabulary in this way, look for any items that characteristically occur at word boundaries, such as specific word endings, particular intonation patterns, and word stress. Stress placement is especially instructive, because most languages place primary word stress on either the initial or the final syllable of every word, or on the second or second-to-last syllable. If your language employs one of these four stress types, it will be relatively easy to identify word boundaries. For example, in a language that stresses the penultimate syllable of every word, you know when you hear a stressed vowel that the end of the word is the following syllable. By the same token, you know that the beginning of the next word immediately follows. Other helpful tricks for getting a handle on word boundaries include asking the informant for help, and preparing for the session in advance by learning as many words and endings as possible.

Segmenting the speech stream into individual sounds is perhaps even harder than segmenting it into words. This is because speech is quantal; in other words, many bits of information are transferred simultaneously. The average human can process up to twenty individual sound blips per second; blips recurring at any frequency greater than this are heard as single sound. Speech sounds, on the other hand, can be processed at rates as high as 40-50 per second (Pinker 1994). If speech sounds were no more than random sound blips, this would not be possible. Linguists have deduced from this fact that humans must transmit and process sounds in chunks, enabling them to pack several speech sounds into each blip of sound that the ear can distinguish. The relevant aspect of this phenomenon for our purposes

is that when humans speak, they do not produce a discrete string of speech sounds—for example, *heat* is not produced as a blip of sound corresponding to [*h*], followed by a blip of sound corresponding to [*i*], followed by a blip of sound corresponding to [*t*]. Rather, the information encoding the sequence [*h i t*] is crammed into a single burst of sound. If you cut this burst of sound into smaller slices, each slice will typically contain information about at least two of these sounds.

What this means for beginning field workers is that it is quite hard to break a stream of speech into individual sounds, unless these sounds are sufficiently similar to ones in their own language, which their ears are already trained to segment into sequences of phonemes. Beginners are often surprised to find that even sound sequences in their own language are difficult to parse when produced out of context. Vowels are notoriously difficult in this regard, but in poor acoustic situations the consonants can be difficult to distinguish as well. For example, it is extremely difficult to tell the difference between English [*s*], [*f*], and [*θ*] over the phone, because phone lines do not transmit the frequencies necessary to differentiate these three sounds. The effects that neighboring consonants have on one another can also add to this confusion. In English, for example, *r*, *l*, and *n* have particularly noticeable effects on preceding vowels. As a result, even native speakers of English can have trouble discerning whether the vowel in *sing* is [*i*] or [*ɪ*], whether the vowel in *thanks* is [*æ*] or [*ɛ*], and so on. Problems like this are of course compounded when one does not speak the language being studied. It can also be difficult to distinguish acoustically similar sounds, such as flapped [*ɾ*] vs. trilled [*r*] vs. [*l*]. Beware, too, of phonetic distinctions that are extremely subtle. Like individual sounds, different features can also be very similar acoustically: the Karabagh dialect of Armenian, for example, has a problematic feature that some people hear as consonant rounding (e.g. *vʷɛn:ə* 'foot'), some hear as velarization (*vʲɛn:ə*), and some hear as a central vocalic onglide (*vəɛn:ə*). Hopefully the phonetic distinctions in the language you choose to work on will not stymie you the way the Karabagh phenomenon (and a very similar one in Ponapean) has done to linguists (including myself), but you should be prepared nonetheless to listen very carefully for difficult distinctions of this type.

The best way for field workers to overcome these challenges is to prepare as much as they can in advance, concentrating in particular on learning the inventory of surface sounds in the language to be studied. Try to prepare a set of words containing each of the sounds in the language, and collect these from the informant. Typically informants will not know or use all of the words in your list, but if this problem arises you can ask them to come up with other words that contain the relevant sounds. Once you have a recording of all of the sounds, listen to them on your own as many times as possible until you feel that you can identify each sound in the language consistently. It is also a good idea to check your transcriptions with the informant, in order to verify that you are hearing the sounds correctly.

One warning: be careful not to fall into the trap of only eliciting words that you already know how to transcribe. Students tend to do this in order to avoid revealing their lack of transcription abilities in front of their peers. If you do this, though, you will never learn to transcribe properly. You *can* learn to transcribe without props in a relatively short period of time, if you are willing to endure a bit of hardship and embarrassment for the first few sessions. Remember—no pain, no gain!

2. The Transcription System

As you are learning how to identify words and sounds, you must also come up with a way of transcribing what you hear. We mentioned in chapter 1 that it is important to use a consistent transcription scheme, with all potentially unclear signs defined in at least one of your

notebooks. Remember that you may know what all of your quirky symbols represent, but others who inherit your notebooks may not, and you yourself may forget!

The transcriptions from which linguists normally choose fall into four categories.

1. The International Phonetic Alphabet (IPA).
2. The official orthographic system already used for the language.
3. An orthographic system based on the dominant language in the area.
4. A standard system already employed by scholars who work on the language.

We strongly recommend using the IPA, because most linguists are familiar with it, and your work will therefore be able to reach the largest possible audience.[3] Using the IPA does have a number of drawbacks, though, which you should bear in mind when deciding which transcription system to use. The set of symbols constituting the IPA (see Appendix) will almost certainly require some modifications to deal with the language you are studying, since it is incomplete in many ways and is not tailored for your language. In addition, it will be largely incomprehensible to your informants, since most of the symbols it employs are not used in standard spelling systems of the world's languages.

In some situations it may be most convenient to choose option (2), the official orthographic system employed for the language you are studying. This option is very favorable for informants, since they will immediately recognize all of your transcriptions if they are literate. Unfortunately, this option is also extremely bad for linguistic purposes, since official orthographies generally neglect to represent phonetic nuances of pronunciation, and can be inconsistent or misleading in their representation of morphophonemic and phonological alternations. Russian orthography, for example, fails to represent the fact that word-final voiced obstruents are pronounced as voiceless. Someone reading the spelling-based transcription <*sad*> for 'garden' could therefore mistakenly conclude that the word was pronounced [*sad*], and would miss the devoicing rule that produces the correct pronunciation, [*sat*]. It is especially unwise to use the official writing system for languages that employ syllabic and logographic systems, such as Japanese and Chinese, because these typically omit even more phonetic details. In sum, we only recommend option (2) for syntacticians, who as a rule are not concerned with phonetic details, and for languages with phonemic writing systems. (In the latter case, make sure that the orthography actually has a one-to-one correspondence between sounds and symbols; if it does not, try to modify the writing system slightly until this is achieved.)

The third option is to use an orthographic system based on the dominant language of the area. When working with Homshetsma, for example, I found it useful to develop a transcription system based on Turkish orthography. This was the best option for several reasons: the phonemic inventories in the two languages are essentially identical; the Turkish system is pleasantly phonemic (i.e. it does not use arbitrary spellings like English does); the characters are easily rendered on a typewriter (most people in northeastern Turkey do not have computers); and the informants can easily read my transcriptions and use the system to write the language on their own, since all of them know how to write Turkish. When I take field notes on Homshetsma for myself, though, I use the IPA, since the Turkish orthography again misses certain phonetic nuances of Homshetsma pronunciation.

The fourth option is to use a standard system already employed by scholars who work on the language. For example, many linguists who work on Armenian use a transcription scheme called the Hübschmann-Meillet system, which differs both from official Armenian

[3] Interested readers can consult the official web site of the IPA at *www.arts.gla.ac.uk/IPA/ipa.html*.

orthography (Armenian has its own script) and from the IPA. Most Armenologists who are not linguists employ the Library of Congress transcription system, which is more closely based on English orthography than on Armenian pronunciation. The fourth option, represented here by the Hübschmann-Meillet and Library of Congress systems, has the advantage of being familiar to other specialists who work on the language, but normally should be avoided, because it will inevitably be difficult and confusing for other linguists and for your informants.

To give you a better idea of the differences between these four strategies, we give here a transcription of the Homshetsma sentence 'the cow that the bear ate was very dangerous' in systems representing each of the four types:

(1)	IPA	*artʰʃona gijadz govɛ ʃad zijonkʰar ɛr*
(2)	Armenian script	*արչոնա գիյած գովէ շատ զիյոնքար էր*
(3)	Turkish-based script	arçona giyadz gove şad ziyonkar er
(4)	Hübschmann-Meillet	*arčona giyaj gove šad ziyonkar er*
	Library of Congress	*archona giyadz gove shad ziyonkar er*

Linguists often like to substitute symbols that are easier to write for symbols that are more accurate. Perhaps the most common example of this is the use of [γ] (voiced velar fricative) for [ʁ] (voiced uvular fricative). Try to avoid doing this, because people who read your transcriptions can easily be deceived into thinking that you are actually transcribing a velar fricative. If you must use substitute symbols, make sure that you indicate in your notes what these symbol actually represent. Do the same for any customized symbols of your own creation; remember that the explanations you provide will be vital for anyone who uses your notes in the future.

When entering field data in a notebook, it is very tempting to use different colors of ink or pencil, as this effectively distinguishes between different types of entries. Though this scheme is esthetically pleasing, we discourage using different colors. One reason for this is that it takes time to switch between pens, and when your informant is spewing out a continuous stream of interesting data you do not want to be wasting your time looking for a different writing implement. The main reason for not using different colors, though, is that the color distinctions will be lost if you photocopy your notes, and the reader will be unable to access the information encoded by your different colors. The need to photocopy notes actually comes up more often than you might think, primarily because you do not want to loan out your original field notebooks to colleagues who are interested in seeing them, lest they be lost. Thus, it is best to prepare for this eventuality by taking notes in one color only, preferably a dark black ink that shows up well on photocopies and does not smudge.

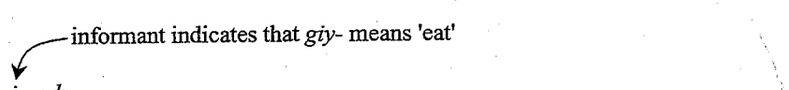
The question then becomes how to convey all of the distinctions that are so easy to make with different colors. We suggest that you try delineating marked information with boxes, underscores, and the like. It is also possible to contrast a thin pen with a noticeably thicker pen, though we do not recommend switching between these two during a session, since it can make you miss valuable pearls of wisdom from the informant. When adding notes at a later date, however, this technique is acceptable.

The last major notation problem involves conveying different types of prosodic boundaries, such as those that divide morphemes, clitics, words, phrases, and sentences. If you do not plan out in advance what symbols you are going to use for each of these boundary

types, you are likely to confuse some of them in your notes in the heat of the moment. We suggest the following symbols:

symbol	**boundary type**
-	clitic
/	morpheme
∪	word (when you originally transcribed as two words what is actually a single word)
I	word boundary missed initially and added later
, or | or ||	phrase
.	sentence

You will also want to indicate whether a given boundary is your postulation, or was pointed out by the informant. It is very important to make this distinction, because for example you may at a later date have trouble remembering whether a given morpheme boundary was your idea or the informant's. Since informants do not normally mention linguistic boundaries of any sort, and when they do they interrupt the elicitation process to do so, we suggest that their suggestions be indicated with a special note nearby, as follows:



Our final notational challenge is sandhi, a process wherein the beginning of one word or morpheme alters the end of the previous word or morpheme, or vice versa. Sandhi rules can make it very difficult to indicate prosodic boundaries clearly. In the Armenian dialect of New Julfa, for example, the future tense prefix *k-* merges with a verb-initial *h-*, yielding the voiced aspirate g^h, as in */k-havadam/* 'I will believe' → [g^h*avadam*]. How is one to represent the morpheme boundary between the *k* and the *h* here? It simply is not possible, unless one presents the underlying sequence of morphemes before it undergoes the merger rule. We suggest that you leave fused forms of this type unparsed, and simply make notations in the margin if necessary.

3. How to Transcribe

3.1. Broad and Narrow Transcription

Even after you have come up with a reasonable set of symbols to use in your transcriptions, there are still many challenges to be faced when transcribing. The most complicated of these involves what are generally referred to as *broad transcription* and *narrow transcription*. Loosely speaking, broad transcription is relatively phonemic; in other words, it represents only the aspects of sounds that are needed to predict the actual pronunciation of words, and to distinguish these sounds from other sounds in the language. Features that do not play a role in distinguishing meaning, such as intonation, are not represented. Narrow transcription, on the other hand, represents more of the phonetic nuances of pronunciation, including some that are not essential in distinguishing meaning. The domain of broad and narrow transcription is obviously a continuum: there can be relatively broad transcriptions, extremely broad transcriptions, relatively narrow transcriptions, and so on, as exemplified in the following representations of my pronunciation of the sentence 'do you want to go to the store?'.

very broad	*du ju want tu go tu ði stor*
broad	*də ju wanə go tə ðə stor*
narrow	*djuw wanə gowɾəðə stoɹ*
very narrow	*djɨɯwãnəgɘɵɾəðəstɵːɹ*

The level of detail that you choose to include in your transcriptions will of course depend on the type of work you are interested in. Syntacticians, for example, will probably prefer fairly broad transcription, since subtle phonetic nuances are rarely relevant syntactically (this is not always true, however—intonation contours and contraction of the *want to* → *wanna* type are syntactically relevant, for example). Broad transcription can also be preferable when your informant is producing vital data and you are not recording the session, and therefore have to transcribe everything on the spot. Some other advantages of using broad transcription are that it is easier to avoid making transcription mistakes on phonetic subtleties, and it is much easier to check your transcriptions with informants, since the transcriptions will be more intelligible to them if devoid of complicated metanotation.

Nevertheless, as a general rule it is better to use a maximally narrow transcription, including as many nuances as you are able, so that you and other linguists who might use your notes will have the most possible information at your fingertips. If you have the wherewithal, you should try to include in your transcriptions not only segmental information such as nasalization, aspiration, length, and so on, but also suprasegmental features such as stress, pitch and intonation, and creaky and breathy voice. Each of these plays a significant linguistic role in languages like English, for example, and you may be surprised at what you find out if you pay attention to features of this type.

When you are transcribing, be very careful to avoid phonemicizing; in other words, avoid the temptation to write down what you know the word "should" be. For example, if someone says [p^h*rabli*] for 'probably', don't transcribe it as [p^h*rabəbli*], the form of the word in careful speech. On the other hand, don't go too far the other way and wildly write down whatever comes into your head. If you think that you may have heard something incorrectly, double check it with the informant or your original recording of the session. You should also be careful about transcribing underlying forms instead of surface forms (an example of this would be writing [p^h*oteto*] for 'potato' when someone actually says [p^h*ət^{h}eɾo*]). It is surprisingly tempting to do this, and you must make a conscious effort to avoid it.

3.2. Transcription Tips

No matter how broad or narrow your transcription is, you should beware of the power of the symbol choices you make. Unfortunately, there is not a one-to-one mapping between sounds and symbols; rather, a given sound can be transcribed in many different ways depending on its linguistic function. For example, a sound that includes labial approximation, rounding, velar approximation, nasality, and voicing could be represented in at least four different ways depending on the context: if the sound is functioning as a vowel, it could be [$\tilde{u}$]; if it is functioning as an approximant, it could be [$\tilde{\beta}^{\gamma w}$] or [$\tilde{w}$]; if it is functioning as a consonant, it could be [$\eta^{\gamma w}$]. Normally the symbol you choose as the basis for your transcription (e.g. *m* in [$\eta^{\gamma w}$]) will be one of the phonemes of the language, and the extra notations will be phonetic details that you have noticed in the pronunciation of this phoneme. If there is no preexisting phoneme list for the language, though, you will have to make your own judgement calls on how to represent each sound.

You can make your task much easier by learning in advance the set of phonemes and allophones in your informants' language. If you can, collect words illustrating each allophone, and try to elicit minimal pairs for each in order to get an idea of how each phoneme can be pronounced. If it is not possible to prepare a list on your own because you cannot find a dictionary or grammar of the language, try asking the informant to come up with a set of words illustrating each sound in the language.

When you are trying to figure out the sounds that your informant is producing, there are a few general tricks you can use to make your life easier. All of these rely on the fact that human languages show certain tendencies, and the language you are working on is very likely to manifest many or most of these. If you can familiarize yourself with these tendencies, then, you can give yourself a head start in figuring out what your informant might be saying. Here are some common cross-linguistic tendencies:

- Voiced consonants (especially obstruents) are often devoiced at the end of a word, especially when not followed by a vowel-initial word.
- Clusters of obstruents tend to assimilate in voicing.
- The final member of a consonant cluster is generally released; non-final members are often unreleased (consider the pronunciation of English *apt*, for example, where the *p* is unreleased and the *t* is released).
- Languages with aspirated consonants often do not aspirate these consonants after *s* (contrast English *tar* [t^h*ar*] and *star* [*star*], for example).
- Nasal consonants tend to take on the place of articulation of a following stop consonant.
- Nasals also cause following stops to become voiced in many languages.

Keeping these tendencies in mind may well prove useful, but be careful not to force your expectations on the data, because a given tendency may not appear in the language you're studying.

4. Transcription Traps

We conclude this chapter with a discussion of some common transcription traps into which both tyros and seasoned linguists often fall. The most common trap involves forcing the phonetic categories of your own language onto the language that you are trying to transcribe. English-speaking linguists, for example, often commit the following blunders:

- *β* (a bilabial fricative) is misheard as *v* (a labiodental fricative).
- *ç* (a palatal fricative, the German <*ch*> or *ich-laut*) is misheard as *ʃ* (a postalveolar fricative).
- *c* (a palatal stop) is misheard as *k* (a velar stop), *k*^{j} (a palatalized velar stop), or *t͡ʃ* (an alveopalatal affricate).
- *t̪* (a dental stop) is misheard as *t* (an apical alveolar stop).
- Unaspirated voiceless consonants are misheard as voiced, and aspirated consonants are heard as unaspirated.
- *r̝* (a voiceless *r*) is misheard as *ʒ* (a voiced palato-alveolar fricative) or *ʃ* (a voiceless palato-alveolar fricative).

The same language transfer phenomenon that causes all of the above mistakes is also responsible for most features of foreign accents; part of the accent of French speakers trying to speak English, for example, results from the fact that they do not aspirate their stops, which is correct for French but incorrect for English.

Even when your language has the same sounds as the target language, it can still be very difficult to transcribe accurately what you are hearing. For example, Japanese */r/* is pronounced in a manner extremely close to the English flap [ɾ], but Japanese speakers typically both hear and produce our flap as a [*d*] or a [*t*] rather than their *r*. This is in part due to their knowledge of English spelling, of course, but there are many other examples where spelling cannot be the culprit. Korean, for example, has both an *l* and an *r*, but Korean speakers nonetheless have great difficulty distinguishing English *l* and *r*. In this case, the problem appears to stem from the fact that *l* and *r* are positional allophones in Korean, whereas they are separate phonemes in English. The liquids *l* and *r* can cause problems even when they are separate phonemes in both the informant's language and the field worker's language: I often mishear Armenian *l* as *r*, and vice versa, even though I have both sounds in my native language, English. The problem in this case presumably results from the fact that the Armenian liquids are pronounced very differently from the English liquids.

Many linguists also have difficulty distinguishing tense-lax vowel pairs, such as *e* vs. *ɛ* and *i* vs. *ɪ*. Speakers of languages that do not make these distinctions typically hear both *e* and *ɛ* as [*e*], both *i* and *ɪ* as [*i*], and so on, or else they hear *e* as [*eː*] and *ɛ* as [*e*], etc. Similarly, English speakers have difficulty distinguishing these pairs because we reinforce the contrast in our own language with length distinctions. The English vowel pair *e* : *ɛ*, for example, as in the words *sate* vs. *set*, is actually realized by most speakers as [*eː(j)*] vs. [*ɛ*]. For this reason, speakers of English typically have a devil of a time distinguishing short [*e*] from [*ɛ*] in other languages.

Another tricky phonetic contrast to beware of involves tones. Some languages have phonemic tone contrast, and others have distinctive prosodic tonal contours. Both of these types of tonal contrasts are very difficult to discriminate consistently for speakers of non-tonal languages such as English. For example, in one of his first meetings with a speaker of Igbo, a colleague transcribed three separate words as homophones. Upon checking these words with the informant, though, he discovered that they were not actually homophones, but rather were distinguished by three different tonal contours. The informant, by the way, was stunned that the linguist had not noticed this "obvious" difference between the three words. The best way to cope with tones is to practice; if you cannot practice with your informants or with recordings of their language, try at least to acquire recordings of another tone language, and familiarize yourself with their tonal contrasts.

One final transcription trap that you should be on the lookout for involves the distinction between consonantal and vocalic features, which we alluded to at the beginning of this chapter. Given that adjacent consonants and vowels typically exert an influence on each other, it can be very difficult to tease them apart and decide which acoustic features belong to which sounds. In the Karabagh dialect of Armenian, for example, it is extremely difficult to tell the difference between a sequence of a plain consonant followed by a front vowel (e.g. *kö*), and a sequence of a palatalized consonant followed by a back vowel (e.g. *k*^{j}*o*). Both are pronounced as something like [*k*^{j}*ö*], but the informant insists (and we agree) that the two are slightly different. There are no hard and fast rules for resolving confusions of this sort, but it can help to ask the informant's opinion, and to examine recordings of the relevant sound sequences with sound analysis software on a computer (for more details, see chapter 6).

The tips and warnings we have dispensed in this chapter should greatly alleviate your initial suffering, if all goes well. Remember, though, that it is impossible to foresee every challenge that will arise, and even when you know what the challenges are, it is not always possible to produce a perfect solution in advance. It is therefore important for you to enter into each

session with an open mind, and a willingness to adapt to the whims and vagaries of the moment.

Suggested Readings

- Voegelin, Charles. 1960. *Guide for Transcribing Unwritten Languages in Field Work.* Manuscript, Harvard University.
- Kelly, John and John Local. 1989. *Doing phonology: observing, recording, interpreting.* Manchester: Manchester University Press.

Exercises

1. Record a brief passage from your informant. Transcribe it broadly and narrowly, indicating as much detail as possible in the latter. Discuss any differences that appear between your two transcriptions.

2. Convert the following excerpt from an article on Robert Mitchum and David Caruso in *Spy* magazine from the IPA into conventional English orthography.

rabərt mɪtʃəm, ðə haliwʊd lɛʤənd, ðə grejtəst æktər ʌv hɪz ʤɛnəreɪʃən, ɪz dɛd. now lɔŋgər wɪθ ʌs. dejvəd kʰəruwsow, ðə formər ɛn waj pʰij dij bluw star, huwz halijwʊd kʰərir tʰʊk ɔf lajk ə rakət dərɛktlij ɪntuw ðə sajd ʌv ə mawntən, ɪz mirlij ʂowvərʂ. kʰamən sɛns wʊd səgʤɛst ðæt ðə flejm-herd dejvəd kʰəruwsow, wʌɾɛvər bægəʤ hij majt bij kʰɛrijiŋ frʌm hɪz kʰərirz kʰəlæps ænd hɪz owvər-pʰʌbləsajzd dəsent ɪntuw ælkəhalɪzəm, wʊd hæv æn ijzijər tʰajm gejniŋ vij aj pʰij æksɛs tʰuw ə bar or rɛstərant ðæn ðə lejt rabərt mɪtʃəm. or wʊd ɪt? æz part ʌv ar ɔngoyŋ ɪnvɛstəgejʃən ɪntuw ðə pʰrəsajs kʰampəzɪʃən ʌv səlɛbrəɾij ʂhatnəsʂ or ʂbʌzʂ, wij ræŋ ərawnd mænhæˀtʰn̩ tʰuw sij wɪtʃ wʌn—ðə dɛd wʌn or ðə ʂfɪnəʃdʂ wʌn—ɪz kʰərəntlij hɛftiŋ ðə mowst ʂpʰəlʂ.

rawnd wʌn: pʰrəmir najt spats ʂʤɛt lawnʤʂ and ʂʤɛt lawnʤ ijstʂ
rabərt mɪtʃəm

ʌs:	*haj. rabərt mɪtʃəm wants tʰuw bij pʰʊt ɔn jər lɪst for tʰənajt.*
ʤɛt lawnʤ ijst:	*jər sort ʌv læst mɪnət, bʌt . . . haw mɛnij?*
ʌs:	*rabərt ænd for gɛsts.*
ʤɛt lawnʤ ijst:	*ʌm, owkej, ajəl pʰʊt ɪt dawn. mɪtʃəm pʰlʌs for.*

dejvəd kʰəruwsow

ʌs:	*hɛlow, aj wantəd tʰuw gɛt mɪstər kʰəruwsow ɔn ðə lɪst. ajm hɪz pʰərsənəl əsɪstənt.*
ʤɛt lawnʤ:	*now, ðə lɪst ɪz fəl for ðɪs ijvniŋ. ɪts tʰuw lejt.*
ʌs:	*ðɪs ɪz dejvəd kʰəruwsow, ðə æktər, ðæt wir tʰɔkiŋ əbawt.*
ʤɛt lawnʤ:	*ajm sarij, ðə lɪst ɪz kʰlowzd.*
ʌs:	*wʌt ɪf ɪts hɪm əlown?*
ʤɛt lawnʤ:	*now. nat ijvən ðɛn.*

rawnd tʰuw: kʰansərt baj ðə arɾəst formərlij nown æz pʰrɪns
rabərt mɪtʃəm

ʌs:	*haj, ðɪs ɪz rabərt mɪtʃəmz pʰərsənəl əsɪstənt. mɪstər mɪtʃəm wʊd lʌv tʰuw gow tʰuw ðə ʃow tʰənajt.*
baks ɔfəs:	*ðə pʰrɪns ʃow? tʰənajt?*
ʌs:	*jæ, hijz rɪlij kʰwajt ə fæn.*
baks ɔfəs:	*ʌm, owkej. haw mɛnij tʰɪkəts dʌz hij want?*
ʌs:	*tʰuw.*
baks ɔfəs:	*rabərt mɪtʃəm. owkej, grejt! ajəl gɛt hɪm sʌmθiŋ gʊd!*

dejvəd kʰəruwsow

ʌs:	*haj, ðɪs ɪz mɪstər dejvəd kʰəruwsowz pʰərsənəl əsɪstənt. aj wəz howpiŋ wij majt bij ejbəl tʰuw skwijz mɪstər kʰəruwsow ɪntuw ðə pʰrɪns ʃow tʰənajt.*
baks ɔfəs:	*wɛl, jər kʰaliŋ pʰrɪɾɪj lejt, juw now. aj mijn, hij kʰæn ʤʌst baj wʌn æt ðə dor.*
ʌs:	*juwʒəlij mɪstər kʰəruwsow dʌzənt hæf tʰuw duw ðæt sort ʌv θiŋ.*