LANGUAGE AND LINGUISTICS 6.2:319-359, 2005
2005-0-006-002-000049-1

Phonological Production in Taiwan Sign Language*

James Myers, Hsin-hsien Lee, and Jane Tsay
National Chung Cheng University

This paper describes an experiment on the production of handshape change in Taiwan Sign Language using the implicit priming experimental paradigm (Meyer 1990, 1991). The results not only provide new evidence that phonological form plays an important role in sign production, but also that the time course of sign production closely matches that predicted by a prominent model of spoken word production (Levelt et al. 1999). The experiment further highlights important methodological considerations in the study of phonological production, not only for sign language, but for spoken language as well.

Key words: sign language, Taiwan Sign Language, phonology, psycholinguistics

1. Introduction

A key function of language is to transmit mental representations through a physical medium, and the role of phonology is to perform the translations closest to the border between the mental and the physical. For functional reasons, then, sign languages require phonology just as much as spoken languages do.[1] Research on sign language phonology has in fact flourished over the past few decades (see overviews in Klima & Bellugi 1979, Padden & Perlmutter 1987, Liddell & Johnson 1989, Coulter 1993, van

* Thanks to the signers who participated, Ku Yu-Shan for signing in the illustrations in this paper, David Corina & Chiu Jung-Shuang for help choosing the response time measure, Chen Jenn-yeu for advice on the implicit priming paradigm, Wang Wenling for help designing the experimental program, Tsao Hsiu-chien for help running the experiment, and the audience at the International Symposium on Taiwan Sign Language Linguistics for comments. I am also indebted to feedback from the two reviewers, in particular David Corina, who did not wish to remain anonymous. The research was funded by National Science Council grant NSC 91-2411-H-194-002.

1 Stokoe (1960) introduced the term *cherology* (*cher* = hand) for this interface system in sign languages, but linguists eventually realized there was no need to be tied to the etymology of the term *phonology* ("study of sound"), any more than etymology determines the synchronic use of linguistic nomenclature like *morphology* ("study of shape") and *syntax* ("arranging together").

der Hulst & Mills 1996, Lucas & Valli 2000, Sandler 2000, among numerous other books, journal articles, and dissertations). This research has demonstrated that in addition to functional similarities, sign language phonology also shares essential formal properties with spoken language phonology, revealing that all human languages involve an "abstract system underlying the selection and use of minimally contrastive units" (Corina 1990:27). In sign languages, these contrastive units include handshapes, which behave like phonemes or distinctive features (as first demonstrated by Stokoe 1960 and Stokoe et al. 1965 for American Sign Language [ASL]). Thus a pair of words may be distinguished solely by the fact that one involves making a fist while the other involves making an open handshape, and each of these handshapes may appear in many other words unrelated in meaning (i.e., the handshapes are phonological rather than morphological units). Moreover, just as in spoken languages, the arrangement of units is also phonologically important in sign languages. Thus a sequence of different handshapes may appear within a single word (e.g., fist to open, or open to fist), and also like spoken languages, not all logically possible arrangements are grammatical (Sandler 1989, 1990, Brentari 1990, 1998, Corina 1990, 1993, Uyechi 1996).

The functional and formal similarities between spoken and sign language phonologies suggest that they may be processed in similar ways as well. When preparing to produce a word, for example, signers should mentally activate similar types of phonological representations and carry out similar operations, in similar orders, as do speakers of languages like Mandarin or English. At a bare minimum, word production in sign languages should involve the activation of some aspect of phonological form, as has been well established from research on spoken languages, both from natural slips of the tongue and speeded reaction time experiments. To date, however, evidence for the use of phonological form in sign production has been somewhat inconclusive. There is no doubt that phonological form plays a role in language errors (so-called "slips-of-the-hand"), as shown by several studies, beginning with Klima & Bellugi (1979). Yet a speeded reaction time experiment reported by Corina & Hildebrandt (2002) failed to show clear effects of phonological form in ASL, raising the possibility, as these authors suggest, that modality differences between spoken and signed languages result in deep differences in phonological processing.

This paper addresses the question of phonological production in sign language with fresh evidence and analyses. The heart of the paper is the description of an experiment on the production of handshape change in Taiwan Sign Language (TSL). We apply the implicit priming experimental paradigm developed by Meyer (1990, 1991) for the study of spoken language but which we use for the first time, as far as we are aware, in the study of a sign language. Our results provide evidence that phonological form does indeed play a role in sign production, though as in the experiment reported by

Corina & Hildebrandt (2002), phonological forms did not affect reaction time directly. However, examining the results within the model of word production presented in Levelt et al. (1999), we argue that the lack of reaction time effects is due to the experimental methodology, a conclusion that paradoxically has quite promising implications. Not only do the overall reaction times and pattern of error rates that we found imply that phonological production in sign language works in a fashion entirely parallel, even down to specific temporal detail, as that in spoken language, but in addition, analysis of our results suggests that our methods may allow researchers to use the study of sign language to illuminate aspects of phonological production that are more difficult to study in spoken language.

Before we describe the experiment and its results, we first provide some background on TSL phonology (§2) and on the study of phonological production (§3). Descriptions of the experiment (§4) and its interpretation (§5) are then followed by a general discussion of its implications for research on phonological production in both spoken and sign languages (§6).

2. Handshape in Taiwan Sign Language phonology

As with all sign languages that have yet been studied, in TSL the form of signs can be analyzed into a relatively small inventory of basic handshapes (see lists given in Smith & Ting 1979, 1984).[2] Also like other sign languages, a subset of these handshapes appears in sign-internal handshape changes. Here we simply illustrate these two empirical observations using our experimental materials as examples, addressing theoretical implications only in so far as they are relevant to our experiment.

Consider first the handshapes described in Table 1 (signs containing these handshapes are illustrated in Appendix A). English names for signs here and throughout the paper come from the TSL primers Smith & Ting (1979, 1984), and following the standard convention in sign language research, names for signs are given in all capitals, to indicate that they are merely convenient labels rather than glosses. The names for the handshapes are English translations for the Chinese names given in Smith & Ting (1979, 1984), which are taken from TSL words in which they prominently appear. Note that these signs are not identical to the eponymous handshapes, since the phonological forms of actual words also require specification of location, orientation, and movement, and in addition may involve both hands, handshape change, and/or non-manual (e.g., facial) features (see Smith 1989 for discussion of distinctive features that can be used to ana-

[2] An updated list of TSL handshapes is also given in the appendix to Chang, Su, & Tai (this volume).

lyze TSL more fully). To avoid confusion with names of actual signs, names of handshapes are italicized and placed in square brackets.

Table 1: Some handshapes used in TSL[3]

Handshape name	Description	Example signs in Appendix A
[*ZERO*]	loosely closed fist, finger tips touching thumb tip	ZERO, CUT CLASS
[*HAND*]	flat open hand, fingers together	HAVE, WRITE, STICK, CUT CLASS
[*SIX*]	thumb and index extended, rest of fingers closed	SIX, FAST
[*SAME*]	open hand with curved, spread fingers	PLACE, RENT
[*LÜ*]	thumb tip touches index tip, rest of fingers closed	LÜ, RICE, WRITE
[*ONE*]	index extended, rest of fingers closed	RICE
[*RENT*]	thumb tip touches middle finger tip, rest of fingers open	RENT, STICK

The phonological status of these handshapes is established by two related arguments. First, there are minimal pairs of words (signs) that are distinguished by use of these handshapes, such as ZERO and HAVE ([*ZERO*] vs. [*HAND*]). Second, pairs of morphologically and semantically unrelated signs can be analyzed as containing the same handshapes (i.e., location, movement, orientation, and non-manual features may differ while handshape does not). For example, as is illustrated in Appendix A, the sign FAST is made with the [*SIX*] handshape, but differs from the sign SIX in orientation and movement. Similarly, the two-handed sign CUT CLASS involves both [*ZERO*] and [*HAND*] on different hands, the sign RICE involves [*LÜ*] on the non-dominant hand (e.g., the left hand for a right-handed signer), the sign WRITE involves [*LÜ*] on the dominant hand, and in the sign STICK, [*HAND*] and [*RENT*] appear both simultaneously (on opposite hands) and sequentially (on the dominant hand).

For readers less familiar with the sign language literature, the analytic decomposability of these signs into handshapes is perhaps less salient than their high degree of iconicity. This typical characteristic of sign languages is actually much less relevant to phonological research than one might think. In addition to the fact that the meanings of signs are more often merely “translucent” from their forms than truly “transparent” (a

[3] LÜ is a family name, and the sign for it mimics the shape of the Chinese character.

distinction made by Klima & Belugi 1979), it cannot be the case that signers derive all aspects of the physical form of signs from meaning directly. For example, such a hypothesis would not explain why the shape of the dominant hand in WRITE deviates from the actual shape of a hand holding a pen or pencil (or chalk or brush, for that matter), nor why this handshape appears in precisely the same form in the semantically unrelated words RICE and LÜ, nor indeed why TSL signers represent the word meaning "write" in anything like this form at all, while signers of other sign languages may use some other form. Regardless of the functional role of iconicity, therefore, a formal theory of phonology is still necessary (for various opinions on the role of iconicity in sign language, see Klima & Bellugi 1979, Armstrong et al. 1995, Taub 2000).

As with spoken language phonology, analyzing signs into basic phonological units quickly leads to important but difficult questions about the structure of the phonological system. In the sign RENT, for example, the final handshape appears to be physically similar to the [*SAME*] handshape seen in the sign PLACE. Thus one analysis would be to consider [*SAME*] here to be lexically specified just like the initial handshape [*RENT*], similar to how spoken languages form words from lexically specified combinations of consonants and vowels. This is the position taken, within very different formal frameworks, by Liddell (1990) and Uyechi (1996). However, most sign phonologists believe that this analysis misses the high degree of predictability between handshapes in the vast majority of sign-internal handshape change: aside from a small set of exceptions (including monomorphemic signs historically derived from compounds), change always involves either all fingers or a specific subset of adjacent fingers, it always involves opening and closing all of the specified fingers (never opening some and closing others), and often, as in the case of [*RENT*] and [*SAME*] in the sign RENT, one handshape is simpler than the other (e.g., all open in the case of [SAME]). This suggests that in some sense one handshape is derived from the other. While differing in technical details and the precise scope of their empirical predictions, Sandler (1989), Brentari (1990, 1998) and Corina (1990, 1993) all present formal phonological analyses of ASL handshape change that capture this key insight. There is yet a third possibility, though, namely that handshape changes should not be analyzed as sequences at all, but rather as wholes related only indirectly to their apparent components, similar to the way affricates are sometimes analyzed (see e.g., Lombardi 1990 and Steriade 1993); Channon (2002) takes a position similar to this in analyses of ASL and other sign languages.

One of our ultimate goals in investigating TSL phonology experimentally was to provide a new source of data to address theoretical questions like these. However, for the purpose of providing background to the experiment described in this paper, which merely attempts to establish that phonological form does indeed play a role in sign production, it suffices to show that handshape change is a genuine aspect of the phonology of TSL. This can be seen quite clearly from the nine signs used as our experimental items, described in Table 2 and illustrated in Appendix B.

Table 2: Target items in the experiment

	Sets	Heterogeneous groupings		
Homogeneous groupings	1. [*ZERO*] > [*SAME*]	FLOWER	SUN	NEW
	2. [*LÜ*] > [*SIX*]	SMART	BEAN	WAKE UP
	3. [*RENT*] > [*SAME*]	NO BIG DEAL	INVENT	NEVER BE-FORE

These signs can be put into three phonologically homogeneous groupings (using the terminology established by Meyer 1990, 1991) so that all three members of a group share the same handshape changes. In Table 2, these groupings are arranged horizontally, with the shared handshape changes described in the Set column (">" represents "changes into"). Note that they differ in most other phonological features (location, orientation, path of movement, and sometimes non-manual features). Following the design established in Meyer (1990, 1991), the nine signs can also be put into heterogeneous groupings (arranged vertically in Table 2) such that the three members in each grouping do not share initial handshapes or overall handshape change. Since all nine signs are monomorphemic, share no obvious semantic features, and represent a variety of syntactic classes, it seems that any possible difference in the processing of the homogeneous groupings versus the hetereogeneous groupings would have to be ascribed to their form.

It should be noted that we assume that the relevant form level here is phonological (representable in terms of abstract categorical features) rather than merely phonetic (involving physical similarities that do not necessarily correspond to abstract features). It is notoriously difficult to separate these levels in practice (see also footnote 4 below), although in the next section we shall review some arguments given by Meyer (1990) and elsewhere for supposing that the experimental paradigm we shall apply does indeed tap into an abstract phonological level. Our target item set inadvertently may help provide another argument, since as pointed out by David Corina (p.c., May 3, 2004), one of our sets may actually involve phonetically similar but not truly phonologically identical target items. Namely, in Set 2, all three target items begin with the thumb and index finger forming a closed ring, but the nature of the finger contact is not the same: two target items (SMART and BEAN) begin with what Liddell & Johnson (1989) call "finger restrained contact" (the index finger nail contacts the thumb pad, ready to be "flicked" off), while the third (WAKE UP) begins with what they term "thumb pad contact". Since this difference is not predictable, it should be treated as phonemic. If our experimental paradigm is sensitive to phonological representations, the Set 2 items, in spite of a great deal of phonetic similarity, should not behave as "homogeneously" as the other two sets, the target items in which do indeed appear to involve precisely the same phonological handshape changes.

3. Phonological production

Phonological knowledge has empirically observable effects not only in the patterns of distributions and alternations studied by linguists, but also in the physical forms analyzed by phoneticians, and in the behavior of language users when perceiving, recognizing, judging, or producing phonological forms. This paper focuses just on one of these sources of evidence, the production of words in isolation. One reason for this focus is the existence of a highly sophisticated model of word production developed by Willem Levelt and colleagues (see Levelt et al. 1999). Armed with such a detailed model, experimental phonology is able to go beyond the traditional search for mere "psychological reality" for various linguistic claims and instead see language use as consisting of processes that occur in real time. Our experiment was designed to begin the investigation into the time course of sign production by using an experimental paradigm also prominent in the development of Levelt's model. In this section we briefly review the relevant aspects of this model and the evidence that has been used to support it (§3.1). We then describe the few studies that have looked at word production in sign languages, and discuss their implications for modeling (§3.2).

3.1 Phonological production in spoken language

The model presented in Levelt et al. (1999) aims to be a complete model of word production in spoken language, and as such describes not only phonological production but also the processes involved when speakers choose words from among semantic competitors, as well as the processing of syntactic features and morphological structure. For our purposes the model can be described as dividing word production into three major stages: stage 1 involves the processing of word information prior to access of phonological form from the lexicon; stage 2 involves phonological encoding; and stage 3 involves phonetic preparation prior to articulation.

This division and ordering should seem quite familiar, since it is quite close to the traditional linguistic view.[4] What makes the model so powerful, however, is the range

[4] Its division of phonology and phonetics into separate stages does not seem to fit well with the "emergent categoricality" approaches presented in Kirchner (1997), Boersma (1998), Steriade

and quantitative detail of empirical evidence that Levelt and his team have collected in support of it, and its consequent degree of precision. For example, experimental evidence has gone beyond previous models in suggesting that stage 2 itself consists of at least two distinct processes: accessing the phonological form from memory (what we shall call stage 2a), and mapping phonological content (e.g., phonemes) into prosodic structure (stage 2b). The research team has even managed to determine estimates for the temporal durations of each stage (Levelt et al. 1998, Levelt & Indefrey 2000). Table 3 gives estimates for these stages in picture naming in milliseconds (msec).

Table 3: Estimated time course for picture naming in spoken language

Stage	Duration	Cumulative time
1. Processing prior to phonological access	275 msec	275 msec
2. Phonological encoding: 2a. Initial access from lexicon 2b. Mapping of units into prosody	125 msec	400 msec
3. Phonetic preparation	200 msec	600 msec

Evidence for these stages, their ordering, and their duration come from a wide variety of sources (summarized in Levelt et al. 1998 and Levelt et al. 1999). The evidence most familiar to linguists comes from natural speech errors (e.g., Fromkin 1971, 1973, 1980, Cutler 1982, Garrett 1980, 1988, Stemberger 1983). Among other things, non-phonological errors tend to operate independently of phonological errors (stage 1 before stage 2), and phoneme deletions, insertions, perseverations, anticipations, and exchanges trigger the application of allophonic processes (stage 2 before stage 3). However, to test more detailed hypotheses about the stages and their time course, experimental methodologies must be used.

One powerful piece of evidence that stage 1 is prior to stages 2 and 3 comes from the picture/word interference paradigm. In this paradigm, pioneered by Schriefers et al. (1990), experimental participants must produce the name of pictured objects while hearing auditory distracters at the same time. Semantically related distracters only affect production latencies (i.e., the duration between presentation of the visual prompt and initiation of articulation) when presented early, while phonologically related distracters only affect production latencies when presented late.

Experimental evidence is also crucial in establishing the distinction between stage 2a (accessing phonological form) and stage 2b (mapping units into prosodic structure).

(2000), Myers & Tsay (2003), and elsewhere. Discussion of such issues, however, goes far beyond the scope of this paper.

One key difference between these stages is that access in stage 2a can be affected by activation of any part of the phonological form, but the mapping process of stage 2b proceeds strictly from left to right (i.e., from the beginning of the word). As pointed out by Levelt et al. (1999), a particularly striking argument in favor of these claims comes from the different effects of explicit versus implicit phonological priming. When primes are presented explicitly, as for example as distracters in a picture/word interference experiment, production latencies for words will be sped up whether the primes match the beginning or ending of the target word (Meyer & Schriefers 1991). Since processing of explicit primes involves auditory access as well as production, the effect here presumably occurs during selection of the phonological form of the target, not during the mapping stage.

The effect of implicit priming is quite different. In the implicit priming paradigm, pioneered by Meyer (1990, 1991), experimental participants are asked to memorize small collections of cue-target pairs; the pairs are designed to ease retrieval of the target without making it entirely predictable (e.g., *house-room*, *bridge-poker*). In homogeneous groupings, the targets are all phonologically similar, while in heterogeneous groupings, they are not. The participants are then presented with the cue words and must produce the associated targets as quickly as possible. The implicit priming effect is defined as a shorter production latency for a given target word when trained in a homogeneous grouping than when trained in a heterogeneous grouping. The assumption is that this effect is due to the implicit primes assisting in the on-line encoding of phonological forms. The alternative possibility that the implicit priming effect is due to mere phonetic factors, such as motor preparation, is rejected by Meyer (1990) and later work because the effect is greater with a greater amount of overlap in the primes, which requires the involvement of a whole-word representation, not just instructions about how to start it. An alternative hypothesis states that the implicit primes merely aid retrieval of phonological forms from long-term memory (at stage 2a). As pointed out by Meyer (1990), the size of the training sets in the implicit priming experiments is much smaller than the size of training sets that memory studies have found are required to aid lexical retrieval; Cholin et al. (2004) also note that immediate serial recall tasks involving phonologically similar words give rise to slower response times, not faster ones as they do in the implicit priming paradigm. Most importantly, the memory retrieval hypothesis is also inconsistent with the finding that implicit priming only occurs if words are phonologically similar at the beginning, i.e., the first phoneme(s) or first syllable(s); by contrast, as noted above, activation of phonological forms in memory at stage 2a can be triggered by phonological cues anywhere in the word. The left-to-right nature of implicit priming implies that the training sessions in this paradigm do indeed allow speakers to prepare part of the left-to-right mapping into prosodic structure.

Implicit priming experiments provide evidence not only for this mapping process but also for the prior stage when the phonological form is accessed. According to Meyer (1991), this stage reveals itself in implicit priming tasks through error rates: opposite from the facilitation in reaction times, training with homogeneous groupings may induce higher error rates than training with heterogeneous groupings. These opposite patterns can be explained if error rate effects occur at the form selection stage, when similar forms compete for attention, rather than at the mapping stage, when implicit primes should help speakers prepare their productions. Meyer (1991) points out that this hypothesis also explains the independent observation that error rate effects are found even when targets overlap only in later parts of the word, while response latency effects only appear when targets overlap from the beginning of the word.

Experiments also allow for estimates of the actual duration of the various stages. The most basic calculation is of the production latency as a whole (i.e., the duration between presentation of the visual prompt and initiation of articulation). Levelt et al. (1998) point out that 600 msec is a slight overestimate for picture-naming times (their own mean production latency was 538 msec), but it seems quite accurate for response times in implicit priming experiments, regardless of language. Thus the mean response times for Dutch words reported in Meyer (1990, 1991) and for Chinese words reported in Chen et al. (2002) were all around 600 msec. Task differences have a larger effect on overall response times. For example, mean production latencies for the Chinese words read aloud in the masked priming experiments in Chen et al. (2003) were all below 500 msec. Presumably such differences in overall reaction time across task are due to different durations for what we label stage 1, i.e., all the processes that precede access of phonological form. The duration of these pre-phonological processes can be estimated from various behavioral and neurological measures (see Levelt et al. 1998, Levelt & Indefrey 2000).

To estimate the combined duration for stages 2 and 3, Roelofs (1997) reviewed a number of experiments using the picture/word interference paradigm, described earlier. Based on a computational model of a variety of such studies, Roelofs (1997) estimated 265 msec as the time from selection of the word to accessing the syllable, a duration that includes all of stage 2 and part of stage 3. The estimate in Levelt et al. (1998) for the duration of stage 2 alone comes from Wheeldon & Levelt (1995), who asked native Dutch speakers fluent in English to decide if the Dutch translation of a word presented in English contained a given phoneme; this task thus required speakers to encode phonological forms of words without actually producing them. The response times for detecting word-initial phonemes in disyllabic words were approximately 125 msec faster than for word-final phonemes. Levelt et al. (1998) then estimated the 200 msec of stage 3 by subtracting the cumulative time of the previous stages from the total production latency. These estimates were found to be consistent with the time course of brain activation patterns in both a magnetoencephelograph (MEG) study (Levelt et al. 1998) and a meta-analysis of many other brain imaging studies (Levelt & Indefrey 2000).

For reasons that will become clear when we describe our own experiment, it is important to note that in all of these speech production experiments, production latency is measured by means of a voice key, that is, a microphone attached to a computer that triggers a signal the instant any sound is made. Thus what is measured is indeed the total duration of all three mental stages, up to the point when speech physically begins. Unfortunately, this means that it is difficult to separate out the effects that are due to stage 2 from those that are due to stage 3. The only methodology described in the literature that is apparently capable of examining stage 2 separately from stage 3 is the cross-linguistic phoneme detection task of Wheeldon & Levelt (1995), a task of limited usefulness due to its reliance on fluent bilinguals with phonemic awareness developed from familiarity with an alphabetic orthography.

Moreover, it is also important to keep in mind that the production model itself is continually undergoing refinement. For example, by going beyond response time measures and including electrophysiologically measured brain activation patterns as well, Abdel Rahman et al. (2003) have argued for a certain amount of parallel processing in word production; stages do not always follow each other in strictly serial fashion. Their specific findings have little direct effect on the time course estimates given above, however, since their evidence suggests only that semantic feature retrieval may continue even after phonological processing has begun; the ordering of morphosyntactic feature (lemma) retrieval prior to phonology is still “serial discrete” in their model (p.858), and their results say nothing about parallelism within phonological processing itself.

Despite such caveats, the model presented above is by far the most explicit and well-tested available, certainly in the study of word production, if not in the study of phonological processing in general. There is no obvious reason why it should not apply to sign languages as well as spoken languages. If so, sign language production should not only also involve activation of phonological units, but it should also show separate stages of phonological processing that parallel both the order and durations of those found with spoken language.

3.2 Phonological production in sign language

Research on word production in sign languages is naturally far more limited than in spoken languages. To date most of what we know comes from language errors, in studies on ASL (Klima & Bellugi 1979, Newkirk et al. 1980, Whittemore 1987) and German Sign Language (Hohenberger et al. 2002). In addition, Corina & Hildebrandt (2002) describe a series of experiments on phonological processing in ASL, including a picture/word interference production task. Here we briefly summarize the major findings and relate them to the production model described above.

As with spoken languages, phonological errors in sign languages operate rather independently from non-phonological errors at the morphemic or syntactic levels (e.g., morpheme or word substitutions), suggesting that the division between stages 1 and 2 is valid for sign languages as well (for linguistic evidence supportive of the same point, see Padden & Perlmutter 1987). Phonological errors themselves treat the various parameters of sign form as independent units, resulting, for example, in perseverations, anticipations, or exchanges of handshape without altering location, movement, orientation, or non-manual features. Moreover, like speech errors, slips-of-the-hand almost never violate constraints of the phonological system; Klima & Bellugi (1979), for example, found that only five of the 131 errors in their corpus contained "extrasystemic" gestures. This suggests that, like spoken language, sign language production involves both stage 2 (encoding of phonological forms) and stage 3 (preparation of phonetic forms, adjusted to fit the phonological system if errors are made at stage 2).

In fact, in phonological errors there seems to be only one major difference between sign and spoken languages (aside from the modality of the units involved). As noted by Hohenberger et al. (2002) in a study of language errors in German Sign Language, signers are far less likely than users of spoken languages to produce exchange errors, where two units switch location (as in the classic spoonerism *sew you to a sheet*). However, as these researchers demonstrate, this is not due to deep differences in processing but rather only to a very superficial effect of modality: the slower speed of the hands relative to oral articulators gives signers more time to catch and correct errors before the complete exchange can be produced, causing them to be realized as anticipations (analogous to *sew you to a seat*). Interestingly, they also emphasize that this conclusion is consistent with another aspect of Levelt's production model not mentioned earlier. This is the self-monitoring process, whereby language producers monitor the output of stage 2 and/or stage 3 before actual articulation begins so that they can block the articulation of erroneous forms. Hohenberger et al. (2002:138) therefore conclude that "signed and spoken language production is, in principle, the same."

When we turn to the production experiment described in Corina & Hildebrandt

(2002), however, the picture at first seems to be somewhat more complex. In this experiment, native ASL signers and native English speakers participated in parallel picture/word interference tasks. The English task worked precisely the same way as Schriefers et al. (1990) (except that only one timing condition was used, with simultaneous presentation of picture and auditory interference word). In the ASL task, participants simultaneously saw the picture whose name was to be signed overlapped with a semi-transparent video image of a signer producing the interference word. In both tasks the interference word was semantically related, phonologically related, or unrelated to the target word. The results showed that for both groups of participants, semantically related targets slowed responses (according to Levelt's model, this is due to competition during word selection in stage 1). However, while the English participants showed very strong facilitation of production latencies from the phonologically related words (i.e., explicit phonological priming), the ASL participants showed no effect at all relative to the unrelated controls.

Null results are notoriously difficult to interpret, but Corina & Hildebrandt (2002) report a similar lack of strong phonological effects for other experiments on phonological processing in ASL. Thus, phonologically related primes had only a weak effect on response times for word recognition in a lexical decision task, and in a handshape monitoring task, native ASL signers failed to perform much better than late learners. An off-line phonological similarity judgment task (described more fully in Hildebrandt & Corina 2002) even failed to find major differences in performance between native ASL signers and hearing participants with no ASL experience.

While all of these experiments imply that phonological form is relevant to the processing of sign language, Corina & Hildebrandt (2002) themselves interpret the results cautiously, commenting that "the behavioral effects of some phonological form-based properties are difficult to establish" (p.108). They speculate that the visual salience of phonological articulation in sign languages eliminates the need for the complex mental machinery that users of a spoken language require in order to reconstruct articulations from acoustic waveforms (according to the Motor Theory of Speech Perception, Liberman 1996). Thus, they argue, mental representations for phonological units and processes simply do not become as active in the minds of signers as in the minds of speakers of spoken languages.

It may be that this interpretation is overly cautious, however. Form-based priming may indeed be weak in sign perception and recognition tasks for the reasons that Corina & Hildebrandt (2002) suggest.[5] Even granting this, it is not clear why production should be as affected by visual salience as perception may be. Producers of signs are not

[5] However, see Moy (1990) for further experimental evidence of the psychological reality of phonological form in sign processing.

trying to reconstruct articulations from perceived forms, but to articulate them in actual fact. Moreover, whether or not Levelt's model is adopted, sign production must involve access of forms from memory, a process that would be greatly simplified if the signs were treated as combinations of a small set of reusable units. Indeed, as we have just seen, some of the best evidence for the psychological reality of phonological units in sign language comes from production data, in particular slips-of-the-hand. The null result of Corina & Hildebrandt's picture/word interference study could be due to any number of factors unrelated to the role of phonological form in production itself. Perhaps the participants were visually confused by the overlapping images, or perhaps the task attempted to probe for phonological processing before signers had actually reached stage 2.

Another factor to consider when pondering the results of reaction time experiments on sign language production (or indeed on any topic) is the method by which the reaction times were collected. This may seem trivial, but its relevance becomes clear as soon as one thinks about the experiments in the context of a specific processing model, such as Levelt's. The method for the picture/word interference experiment is not described in Corina & Hildebrandt (2002), but according to David Corina (p.c., September 26, 2002), it involved the use of an infrared trip beam to signal the instant when participants raised their hands to begin signing. The timing of this same event can also be measured by the keyboard lift-off method, which Corina has successfully used in a lexical decision experiment on Spanish Sign Language. In this method (which, unlike the trip beam method, requires no special equipment), the experimental participant begins by resting his or her hands on a key on a computer keyboard (e.g., the space bar). When he or she receives a visual prompt on the computer screen, the hands are lifted and the computer records the time between the onset of the visual prompt and the release of the key press.

Crucially, note that either method records the timing of a very different event from that recorded by the voice-key method used in spoken language experiments. For speakers what is measured is the instant when sound is produced by their mouths, which occurs at the end of stage 3. By contrast, for signers what is measured is the instant when they have decided that they know enough about the phonological form to begin signing, which is certainly well before the end of stage 3. In fact, it is likely to be soon after initial contact is made with phonological forms accessed from the lexicon in stage 2a. Since large articulators like the arms are so slow compared to oral articulators, there is plenty of time for stages 2b and 3 to be mentally prepared as the hands are being lifted into signing position. Therefore, differences in results for experiments on spoken vs. sign languages may not be due to deep differences in phonological processing at all, but rather differences in the stage of phonological processing that is probed by the voice-key method vs. the trip-beam or lift-off methods. This hypothesis will be explored more fully later.

4. An implicit priming experiment on TSL

The goals of this experiment were threefold. First and most fundamentally, we simply wanted to know whether it was possible to perform an implicit priming task on a sign language, since it apparently it had never been tried. As we saw above, there are virtually no psycholinguistic studies on sign languages that have used reaction-time measures at all. Second, we wished to test the suggestion made by Corina & Hildebrandt (2002) that phonological form does not play an important role in the online processing of sign language, a suggestion made partly on the basis of a picture/word interference task conducted on ASL. Our experiment was intended to provide data from a new language (TSL) using a new production task (implicit priming). This task was chosen not only for its relative simplicity compared to the picture/word interference task, but also because, as explained in §3.1, it is in principle capable of providing independent information on two stages of production: access (stage 2a) and mapping (stage 2b) of phonological forms. Finally, the experiment was planned as the first of a series examining the time course of phonological encoding in sign production. If this first experiment was successful, in the future we hoped to apply the implicit priming paradigm again, this time using materials that would allow us to test whether phonological units are mapped left to right in sign language. Among other things, determining this should be able to shed light on the phonological nature of handshape change.

4.1 Methods

We followed the procedures for the implicit priming task described in Meyer (1990, 1991) as closely as possible.

4.1.1 Participants

Twenty deaf, fluent TSL signers were paid to participate in this experiment. Nine were female, eleven male, and their ages ranged from 14 to 59 years old (average about 40 years old). All used TSL as their primary language, though all were also able to read and write Mandarin Chinese. Five were also able to speak and lip-read some Mandarin, one some Southern Min, and one a little of both. Only four signers could be classified as "native" according to the strict criterion used by Hildebrandt & Corina (2002) (i.e., they acquired TSL from deaf parents), but the rest were exposed to TSL before the onset of puberty. Thus the age of TSL acquisition ranged only up to 11 years old, with the average being 7 years old. An additional seven TSL signers (including three who learned TSL when already older than 10 years old) were paid to participate in a pilot using dummy materials to test the procedure and reaction time measures, and their results were not analyzed.

4.1.2 Materials, design, and procedure

As described earlier, we followed Meyer (1990, 1991) in choosing our materials so that they could be arranged in two ways, either in groupings of words that were phonologically similar, or in groupings of words whose phonological forms shared nothing in common. In this particular experiment, similarity involved sharing the same handshape change, while location, orientation, movement path, and non-manual features were allowed to vary. We settled on the nine one-handed signs shown earlier in Table 2 (see also Appendix B).

In order to trigger the production of these target items, each was associated with a cue word or phrase, presented visually in Chinese. These cues, with their associated targets, are listed in Table 4 below. The associations were designed merely to assist memorization of the otherwise arbitrary cue-target pairs; the nature of the association was not an experimental variable.

Table 4

Cues (Chinese)	Targets (TSL)
情人節 (Valentine’s Day)	花 FLOWER
熱 (hot)	太陽 SUN
手機 (cell phone)	新 NEW
第一名 (no. 1)	聰明 SMART
貢糖 (candy)	豆 BEAN
起床 (get out of bed)	醒 WAKE UP
討厭 (annoying)	不屑 NO BIG DEAL
科技 (technology)	發明 INVENT
殺人 (murder)	從來沒有 NEVER BEFORE

Note that unlike most experimental paradigms used in lexical research, lexical frequency is ignored in the design of implicit priming experiments, other than ensuring that cues and targets are familiar to all participants (see Meyer 1990). This is partly because response times for individual items depend not only on characteristics of the target, but also on the characteristics of the cues and how they relate to the targets. Since any given target is always preceded by the same cue, there is no way to separate out these effects; they are inherently confounded. This is not a problem, however, since the crucial comparison in this paradigm relates to the effect of context, that is, homogeneous vs. heterogeneous groupings. Thus each item acts as its own control: the only difference between a homogeneous vs. heterogeneous trial is the training context.

Cue-target pairs were trained in either homogeneous groupings (i.e., the horizontal groupings in Table 2) or heterogeneous groupings (i.e., the vertical groupings in Table 2). Specifically, participants were told in TSL (by a hearing but fluent-signing experimenter) which target word should be produced for each written cue. During each of these training phases, the participant practiced until he or she was able to produce the expected target reliably.

After a grouping of cue-target pairs was trained, each participant was presented with a block of nine trials in which production latencies were measured. The block contained the three cues that had just been trained, each repeated three times, with all trials presented in random order but adjusted so that no item appeared two times in a row. The reaction-time phases of the experiment were run on a laptop computer (PC clone running Windows Me), with experimental control handled by E-Prime 1.0 (Schneider et al. 2002). Production latency was measured using the keyboard lift-off method. To begin each trial, participants were asked to place their dominant hand on the keyboard, with their index finger depressing the space bar. The trial then began with the display of the symbol + on the center of the screen, merely to orient the eyes to the correct location, which after one second was replaced by a cue word or phrase. Participants then had to lift his or her hand and begin signing the correct target word as quickly and as accurately as possible, without any hesitation. The computer recorded the time between the onset of the display of the cue word and the release of the space key (i.e., when the hand was lifted). The fluent-signing experimenters then immediately coded responses into four categories: correct, wrong word choice, hands hesitating on keyboard, and hands hesitating in the air after leaving the keyboard.

After each block was completed, the participant would then receive training in another grouping of three cue-target pairs, followed by the relevant cue-production trials on the computer, and so forth until three repetitions of each block were completed. The order of blocks was randomized, with homogeneous and heterogeneous blocks mixed together. The primary purpose of all this repetition (an inherent aspect of the implicit priming paradigm) was to increase the total number of trials so that statistical analysis was possible. Each item appeared equally often in homogeneous groupings as in heterogeneous groupings. Thus each item appeared 18 times during the course of the experiment (2 grouping conditions × 3 repetitions of blocks × 3 repetitions of items within each block), with a total of 162 trials (18 × 9 cue-target pairs).

Participants were arbitrarily assigned to two equal-sized groups, defined by whether or not the first block they were exposed to was a homogeneous or heterogeneous block (following standard procedures, this was done in case the first training experience colored the participant's behavior throughout the rest of the experiment). Each participant required 30 to 40 minutes to complete the experiment.

4.2 Results

We analyzed two aspects of the responses: response times (production latencies) and error rates. Errors consisted of all responses coded as errors by the experimenters during the experiment (i.e., wrong word choices and hesitations), plus responses with latencies of one second or longer (the same criterion used by Meyer 1990). To prepare response time (RT) for analysis, we grouped responses by condition (heterogeneous, homogeneous), and within these, by set (Set 1, Set 2, Set 3), and within these, by repetition of blocks (repetition within blocks was not separated out for analysis). We then calculated the average RT for each combination of condition, set, and repetition. Note that "set" here refers to the set of items in the design (i.e., the items appearing horizontally in Table 2), not necessarily the grouping of items that appeared within a block during the experiment. Thus the words that appeared in the analysis labeled "heterogeneous set 1" were the same as those in "homogeneous set 1". The only difference was that heterogeneous set 1 consisted of responses to words in Set 1 when they were trained and tested along with words of different phonological types (e.g., FLOWER when trained with SMART and NO BIG DEAL), while homogeneous set 1 consisted of responses to words in Set 1 when trained and tested with words sharing handshape change (e.g., FLOWER when trained with SUN and NEW). To prepare error rates for analysis, we calculated the proportion of errors (as defined above) within each combination of condition, set, and repetition.

As required by the design, statistical analyses for both RT and error rates were conducted within participants, but we also noted what order group (heterogeneous first, homogeneous first) that each participant belonged to and included this as a between-participant variable. We then performed separate four-way ANOVAs on RT and error rates (order × condition × set × repetition).[6] Theoretical interest lies primarily in any main effect and interaction involving the conditions and the sets (order and repetition were only included to understand what role, if any, practice had on the responses

[6] One data point was missing in the RT analysis (one participant, in one condition, set, and repetition, made only errors, leaving no mean RT). We estimated this missing value following Winer (1971:488-9).

over the course of the experiment).

The effects of condition and set on response time are illustrated in Figure 1 below; the same information is given in Table 5, along with standard errors.

Figure 1: The effects of condition and set on response time

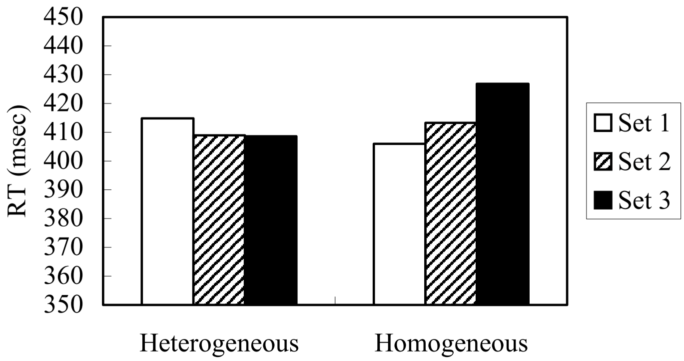


Table 5: Means in msec (and standard errors) for reaction times

	Heterogeneous	Homogeneous
Set 1	415 (11.3)	406 (11.8)
Set 2	409 (11.5)	413 (12.6)
Set 3	409 (10.9)	427 (13.0)

As hinted at by the large standard errors relative to the differences in RT, there was no main effect of condition; mean response times for the heterogeneous condition (411 msec) and homogeneous condition (415 msec) were not significantly different at the 0.05 level ($F(1,18) = 0.4$, $p = 0.53$). There was also no main effect of set; mean response times for Set 1 (410 msec), Set 2 (411 msec) and Set 3 (418 msec) were not significantly different ($F(2,36) = 1.13$, $p = 0.33$). However, there was a significant interaction between condition and set ($F(2,36) = 4.5$, $p = 0.02$), which is reflected in the different pattern of bar lengths in the left versus the right side of Figure 1. In particular, it appears that in the heterogeneous condition, there was very little difference in response times across the sets, while in the homogeneous condition, differences were much more pronounced, with Set 1 the fastest and Set 3 the slowest. No other effects or interactions were significant (all ps > 0.3), implying that there were no effects of practice on RT over the course of the experiment. The lack of a main effect of condition was apparently not due to the influence of a few recalcitrant items, since as shown in Table 6, about half of the items showed longer RTs in the heterogeneous condition, while half showed the opposite tendency.

Table 6: Mean RTs (msec) by item

	Heterogeneous	Homogeneous	Difference
花 FLOWER	420	417	3
太陽 SUN	404	395	9
新 NEW	414	411	3
聰明 SMART	420	404	16
豆 BEAN	402	403	-1
醒 WAKE UP	408	418	-10
不屑 NO BIG DEAL	420	420	0
發明 INVENT	416	427	-11
從來沒有 NEVER BEFORE	404	433	-29

The effects of condition and set on error rates are illustrated in Figure 2 below; means and standard errors are given in Table 7.

Figure 2: The effects of condition and set on error rates

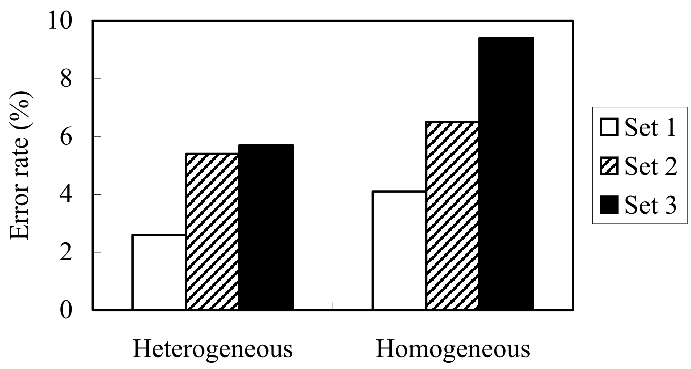


Table 7: Means (and standard errors) for error rates

	Heterogeneous	Homogeneous
Set 1	2.6% (0.8)	4.1% (1.1)
Set 2	5.4% (1.2)	6.5% (1.2)
Set 3	5.7% (1.1)	9.4% (2.0)

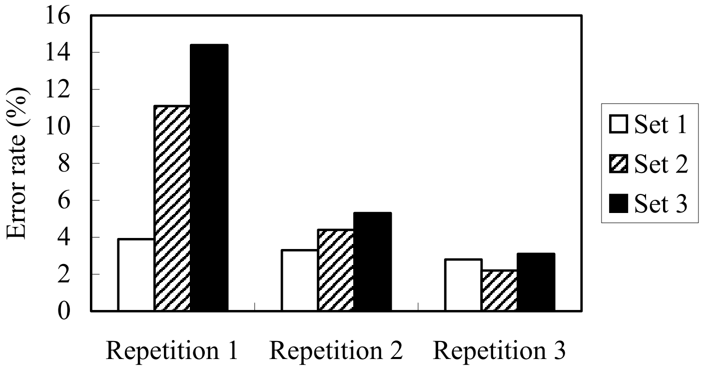
This time there was a main effect of condition, quite a large effect in fact. The mean error rate for the heterogeneous condition (4.6%) was significantly lower than for the homogeneous condition (6.7%) ($F(1,18) = 9.13$, $p = 0.007$). As shown in Table 8, this pattern was consistent, being found in six out of the nine items (with only one item showing the opposite).

Table 8: Mean error rates by item

	Heterogeneous	Homogeneous	Difference
花 FLOWER	2.2%	5.6%	-3.4%
太陽 SUN	2.2%	1.7%	0.5%
新 NEW	3.3%	5.0%	-1.7%
聰明 SMART	7.8%	7.8%	0.0%
豆 BEAN	3.3%	5.0%	-1.7%
醒 WAKE UP	3.3%	6.7%	-3.4%
不屑 NO BIG DEAL	6.1%	13.9%	-7.8%
發明 INVENT	8.9%	8.9%	0.0%
從來沒有 NEVER BEFORE	3.9%	5.6%	-1.7%

There was also a main effect of set, with the mean error rate for Set 1 (3.3%) lower than that for Set 2 (5.9%), which was in turn lower than that for Set 3 (7.6%) ($F(2,36) = 4.0$, $p = 0.03$). However, there was no significant interaction between condition and set ($F(2,36) = 0.75$, $p = 0.48$); unlike the case with the response times, the pattern of increasing error rates from Set 1 to Set 3 was basically the same in both conditions. There were also two significant effects relating to repetition. First, there was a main effect of repetition ($F(2,36) = 27.6$, $p < 0.0001$), with the error rate for repetition 1 (more accurately, the first presentation of the materials) being higher (9.8%) than for repetition 2 (4.4%), which was higher than for repetition 3 (2.7%). This merely shows an effect of practice on reducing error rates. Somewhat more interesting was a significant interaction between repetition and set ($F(4,72) = 5.49$, $p = 0.0006$). This interaction is illustrated in Figure 3, where it can be seen that the difference across sets was mainly found in repetition 1, when participants had their first contact with the materials. After some practice with them, this effect disappeared.

Figure 3: The effects of repetition and set on error rates



Error rates also showed a nearly significant interaction between condition and order ($F(1,18) = 3.82$, $p = 0.066$), since participants who received a homogeneous set first tended to show a larger difference in error rates between the two conditions than did participants who received a heterogeneous set first. No other effects were significant (all *p*s > 0.35), in particular the interaction between condition and repetition: while practice reduced overall error rates and reduced error rate differences across sets, it had no effect on reducing the different patterns of responses to homogeneous versus heterogeneous conditions.

One final issue that should be mentioned before we move into the discussion is the possible role of age of acquisition in the results. Studies (e.g., Mayberry & Fischer 1989, Hildebrandt & Corina 2002) have found differences in the performance of signers born to deaf signers versus those born to hearing parents (who are thus typically not exposed to a sign language until they enter school) in how they perceive phonological forms. Though all of the signers in our experiment acquired TSL prior to puberty, only four of them were, strictly speaking, native signers (i.e., born to deaf signers). Nevertheless, when we looked for evidence that native competence played any role in our results, no such evidence was found: in new ANOVAs for RT and error rates that included native competence as a between-participant factor, this factor showed no main effect and did not interact with any other factor.

5. Discussion

If nothing else, the experiment fulfilled its first goal: we demonstrated that it is possible to run an implicit priming task on a sign language and obtain meaningful results. The most important of these results related to the effect of condition: in both response times and error rates, we found that the difference between heterogeneous and homogeneous conditions had an effect. In response times, this effect was indirect, being found only in a differential patterning across the sets in the two conditions. In error rates, the effect was quite robust, with items in the homogeneous condition being produced with higher error rates (i.e., hesitations both before and after lifting the hands from the keyboard, and production of the wrong word). Thus our experiment has provided evidence for form-based effects on the production of signs. Nevertheless, as with Corina & Hildebrandt's (2002) experiment, we failed to find a main effect of reaction time, with our most robust effects appearing instead in error rates.

Before discussing how this pattern of results should be interpreted, we first examine a factor that did have effects on both reaction times and error rates: set. As shown by error rates (in both conditions) and response times (only in the homogeneous condition), items in Set 1 ([*ZERO*] > [*SAME*] signs) seemed to be easier (lower error rates, faster response times) than items in Set 3 ([*RENT*] > [*SAME*] signs), with items in Set 2 ([*LÜ*] > [*SIX*] signs) falling in between. It is important to resist the temptation to interpret these differences as necessary consequences of the differences in phonological forms of the words in these sets, since the sets also differed in at least three other ways: the lexical frequency or familiarity of the target forms, the lexical frequency or familiarity of the Chinese cues used to prompt the signers, and the associative relations between the cues and the targets. Of these factors, the only two about which we have concrete information are the lexical frequency or familiarity of the cues and targets. Since the prompts were Chinese words or phrases, we can look up their frequencies in a large corpus. In our case we ran searches for them on www.google.com (see Blair et al. 2002 for evidence that Internet search engines provide reliable frequency estimates). The results are shown in Table 9.

Table 9: Estimated frequencies of Chinese cues (www.google.com, 9:30 am 2/21/2003)

Set 1	情人節 (Valentine's Day)	熱 (hot)	手機 (cell phone)	Average
Frequency	119,000	2,330,000	2,240,000	1,563,000
Set 2	第一名 (no. 1)	貢糖 (candy)	起床 (get out of bed)	Average
Frequency	290,000	1,660	632,000	307,887
Set 3	討厭 (annoying)	科技 (technology)	殺人 (murder)	Average
Frequency	299,000	9,670,000	640,000	3,536,333

Although the average for Set 3 ends up being the highest, this is due solely to the unnaturally high frequency of the word meaning "technology", likely reflecting the bias of webmasters more than anything else. Removing this item gives Set 1 cues the highest average frequency and makes the frequencies for Sets 2 and 3 roughly comparable.

We also know something about the lexical frequency or familiarity of the TSL target signs themselves. The TSL textbook series by Smith & Ting (1979, 1984), like any good language textbook, introduces vocabulary in a sequence judged to be the most useful. Thus the division of vocabulary across the two volumes can be taken as a reasonable estimate of vocabulary usefulness, and hence of frequency and familiarity. Applying this to the current experimental materials, we observe that all three of the items in Set 1 are introduced in Volume 1, all three of the items in Set 3 are introduced in Volume 2, and the items in Set 2 are mixed (two are introduced in Volume 1, and one is introduced in Volume 2).

Thus the lower error rates for Set 1 are associated not only with more familiar Chinese cue words, but also more familiar TSL targets, while the higher error rates for Set 3 are associated with a lower degree of familiarity in both Chinese prompts and TSL targets. Another clue that differences in error rates across the sets were due to frequency effects rather than phonology comes from the interaction between repetition and set on error rates, illustrated earlier in Figure 3. This interaction shows that the set difference effect was solely due to participants' first exposure to the items. This is what one would expect from frequency effects, which can be counteracted by repeated exposure. By contrast, the error rate difference between the homogeneous and heterogeneous conditions did not wane during the course of the experiment.

Again we must clarify that in contrast to most experimental paradigms used in research on word processing, frequency effects themselves are not really relevant here, except to show that, unsurprisingly, non-phonological factors played a role in our experiment. A deeper analysis of frequency effects is not possible due to the inextricable confounding between cue and target properties, and in any case, such an analysis would tell us less than one might expect. For example, it may seem, as David Corina (p.c., May 2, 2004) has suggested, that frequency effects, or the lack thereof, could be relevant in determining the processing stage probed in our experiment: only lexical stages of processing should show such effect. However, in Levelt's model all stages are lexical to some degree: even the phonetic encoding stage involves retrieval from a lexical syllabary of stored articulatory gesture programs (Levelt & Wheeldon 1994, Cholin et al. 2004). Thus frequency effects should be ubiquitous in any appropriately designed experiment.

Although non-phonological differences across sets are likely to be the primary factors causing the different error rates, we should also briefly consider the possible influence of the degree of homogeneity within each set. Recall that when we introduced the materials we noted that the target items in Set 2 do not seem to be fully phonologically homogeneous: SMART and BEAN begin with finger restrained contact, while WAKE UP begins with thumb pad contact. The set may thus be an example of an "odd-man-out" set (in the terms of Cholin et al. 2004) and should therefore be expected to show weaker effects than the other two sets, which were fully homogeneous. Though there was no significant interaction between set and condition in error rates, there was indeed a trend in precisely this direction: as can be seen from Table 7 above, Set 2 showed a smaller difference in error rate between the homogeneous and heterogeneous conditions (1.1%) than either Set 1 (1.5%) or Set 3 (3.7%). The lack of significance and the possible influences of non-phonological factors here mean that we should take this observation with a great deal of caution, but it may be worth following up in future studies.

We now turn to a discussion of the most important finding of the experiment: robust error rate effects without main effects of reaction time. As noted earlier, higher error rates in homogeneous contexts are also commonly found in implicit priming experiments performed on spoken languages, and, beginning with Meyer (1991), this has been taken to suggest that error rate effects occur at stage 2a, when phonological forms are first being accessed from memory. Yet unlike most implicit priming experiments conducted on spoken languages, we failed to find differences in overall reaction time across the two grouping conditions, thus missing the effect that has been claimed to occur at stage 2b, when phonological units are being mapped into prosodic structure. Here we consider two factors that may have affected our results, viewed within the framework of Levelt's production model.

The first factor is the phonological structure of our experimental materials. It is possible that the phonological forms in our homogeneous groupings, while indeed similar, were not similar "from left to right". That is, although they shared phonological elements (the handshape change, including the first handshape in the change), they differed in other parameters at the beginning of the sign (in particular, location and orientation). Thus the set of phonological features linked to the initial timing slot in the prosodic structure (i.e., the first "segment") would not have been identical across the items even within a homogeneous grouping. Roelofs (1999) has shown that in Dutch, mere featural similarity in onsets (e.g., /b/ and /p/) was not enough to trigger the implicit priming effect; onset segments had to be identical in all features. Similarly, Chen et al. (2002) found that when Chinese syllables matched only in tone, which like handshape change is distributed across the entire syllable, there was no standard implicit priming effect either. It is thus possible that phonological differences between the signs we used and the forms used in most spoken language implicit priming experiments may have led them to be processed in different ways.

The phonological structure of the materials is certainly an important factor to keep in mind for future experiments. However, we believe that a second, methodological factor may have had a much greater influence in creating the pattern of our results. Namely, our use of the lift-off method to measure response times may mean that we tapped into an earlier stage of word production processing than the voice-key method used for spoken languages. According to the argument sketched in §3.2, the signers in our experiment must have lifted their hands after achieving initial access of phonological forms at stage 2a, before the mapping of stage 2b could even begin. This hypothesis would immediately explain the significantly higher error rate in the homogeneous condition (due to processing at stage 2a) and the lack of RT differences (missed since stage 2b had not yet been reached).

A key further prediction of this hypothesis is that the overall response time in our experiment, missing as it did stages 2b and 3, should be quite a bit faster than those observed in spoken language studies. In fact, with the durations of these stages estimated as in Table 3, we can make this prediction quantitatively precise: our overall reaction time should be around 200 msec faster (an overestimation for the duration of stage 3, i.e., stage 3 plus a bit of stage 2). Recall that in spoken languages, production latencies in implicit priming experiments are around 600 msec, a value that is consistent across word length, language, and size of the cue-target sets (see e.g., Meyer 1990, 1991 for Dutch, Chen et al. 2002 for Chinese). By contrast, as can be seen from Table 5 above, our average response times for TSL were just a little over 400 msec. More precisely, the mean RT over all 360 data points used in the RT analysis (2 conditions × 3 sets × 3 repetitions × 20 participants) was 413 msec (standard deviation 92 msec). Our lower overall RT cannot be due to how we eliminated erroneous responses, since we followed standard methods here as well (e.g., Meyer 1990 also rejected RT values over one second). The difference also cannot be ascribed to differences in manual vs. oral articulation, since in both modalities what is measured in these experiments is the time before articulation actually begins, and in any case manual articulation is slower, yet our response times were faster. These observations suggest that not only were our signers lifting their hands from the keyboard prior to stage 2b, but that the durations for the preceding and following stages were approximately the same as those deduced for the production of spoken languages.

Yet another argument for these conclusions comes from the different RT patterns across sets in the homogeneous vs. heterogeneous conditions. Recall that in the heterogeneous condition, RT values were quite close across Set 1, Set 2, and Set 3, while in the homogeneous condition, Set 1 was fastest while Set 3 was slowest, consistent with the difference in error rates (highest in Set 3 and lowest in Set 1). As argued earlier in this section, the cross-set differences were likely due to frequency effects, not phonological properties. Now, Meyer (1991) noted that the effects due to what we call stage 2a were not only different from those due to stage 2b, but were also more sensitive to varying aspects of the materials, such as "the strength of the associations between prompts and response words, the relative frequencies of the words, and the semantic relations among them", so that stage 2a effects could arise "only if several of these factors conspired in making the selection of the response words particularly difficult" (p.85). Applying this view to our own experiment, we expect to find response time differences across sets to show up more strongly in the homogeneous condition, when items are competing phonologically, since this competition would add to the "conspiracy" that makes word form selection (stage 2a) sufficiently difficult to affect responses. This is in fact just what the interaction between condition and set seems to show.

Summarizing, then, where our results differ from those found with spoken languages, it is apparently primarily because a difference in methodology (lift-off vs. voice-key) led to our probing into an earlier stage of phonological production than the implicit priming experiments that have been conducted on spoken languages. The production process itself seems to be identical across modalities, even down to the detailed time course of the stages.

6. Conclusions

Intensive linguistic research over the past few decades has established beyond any reasonable doubt that sign languages employ phonological systems quite comparable, in both function and form, with spoken languages. In language production in particular, the evidence is quite strong that phonological units like handshapes are manipulated mentally by signers just as producers of spoken languages manipulate phonemes and features. The null results of the ASL production experiment reported in Corina & Hildebrandt (2002) are the sole anomaly in the previous literature, but like all null results, they merely provide a call for further research.

The experiment described in this paper not only provides further evidence for the mental processing of handshape in the production of signs, but also suggests a possible reason for the null results in Corina & Hildebrandt's experiment: the measure of reaction time they used may have tapped into an earlier stage of processing, before all aspects of phonological form were fully fleshed out in signer's minds. This hypothesis is supported by a variety of arguments, including the precise duration of the reaction times. Too often psycholinguists analyze reaction times merely to find out if they are different across conditions, without considering the information provided by the absolute values themselves. After all, reaction times represent the duration of real processes occurring in real time. The analyses presented in this paper demonstrate that an understanding of precisely what is being measured in an experiment can be crucial. In this case, they suggest that sign language processing shares quite deep similarities with spoken language processing, even down to the ordering and fine temporal detail of the stages. The statement of Hohenberger et al. (2002) declaring the production of sign and spoken languages to be the same is even more accurate than they may have realized.

Nevertheless, it must be admitted that something of a methodological challenge is presented by the discovery that the keyboard lift-off method used for sign language taps into a different stage from the voice-key method used for spoken language. Namely, if researchers are interested in the time course of the final stages of sign production, stages apparently missed by the lift-off method, some other method for measuring reaction times must be developed. One possibility would be to use high-speed video and then estimate response times by counting frames. High-speed video would be necessary, since response time differences across conditions (as estimated from research on spoken languages) are too close to the limits of temporal resolution of standard video (about 30 msec). Yet using video to measure response times is not only very labor-intensive, but there are also questions about its inherent reliability. A voice-key is triggered the instant the mouth begins to make noise, but how should one objectively define the precise moment when a signer truly begins to articulate a sign?

Regardless of the method for measuring RT, the implicit priming method may also be somewhat problematic for further study of handshape change in particular. As noted in §2, a primary reason for the interest in handshape change is the question of whether it is represented as a sequence of handshapes or as a whole. Phonological analyses in the sign literature typically assume that in some sense handshape changes are composed of separate (though autosegmentally linked) handshapes (e.g., Sandler 1989, Brentari 1990, 1998, Corina 1990, 1993). The handshape detection experiment described in Corina & Hildebrandt (2002) provided some psycholinguistic evidence for this view, since no RT difference was found between detecting a handshape in a sign without handshape change and detecting the first handshape in a sign with handshape change (detecting the second handshape naturally took slightly longer, since participants had to wait until the sign neared completion). Yet if we want to address the question of handshape change composition with an implicit priming experiment, we face a possible problem with the materials.

Suppose we want to know if production of the sign FLOWER really begins with access of the individual handshape [*ZERO*]. The natural thing to do would be to include FLOWER and ZERO together in the homogeneous training condition. Unfortunately, however, these signs would have a different number of handshapes and thus possibly different prosodic structures. Research on spoken languages has shown that the implicit priming effect only occurs if items in the homogeneous condition are prosodically identical (see e.g., Roelofs & Meyer 1998). It may seem better, then, to train FLOWER along with another handshape-change sign that also begins with the [*ZERO*] handshape, but unfortunately this is impossible. As noted earlier, each of the handshapes that participate in handshape change is typically predictable from the other. Thus handshape changes that begin with [*ZERO*] necessarily end with [*SAME*], just as in FLOWER. Therefore, further research on the production of handshape change seems to require the development of methods beyond those currently described in the literature.

On the other hand (so to speak), what are limitations for our original research questions may ultimately prove beneficial to the study of phonological production in general. The lift-off method appears to tap into a stage prior to articulatory preparation, perhaps prior even to the mapping of phonological units into prosodic structure. As far as we are aware, no method developed for spoken language has this capability. The closest seems to be the method of Wheeldon & Levelt (1995), yet as noted earlier, this method has limited usefulness for many languages. The challenge is that there is no physical correlate (short of data that could only be collected through expensive and time-consuming brain imaging studies) for the moment when a speaker has accessed the phonological form of a word, but has not yet begun to flesh it out. The lift-off method, however, appears to provide just such a physical correlate. Given that all the evidence so far points to the conclusion that language production works precisely the same way for sign and spoken languages, research on sign language could thus provide insights into the working of spoken language that would not be available any other way. This provides yet another argument for the position championed for forty years by sign language researchers: far from being an exotic novelty, sign languages can actually provide crucial insights into the human language faculty that would otherwise never be uncovered.

Appendix A: TSL signs containing the indicated handshapes

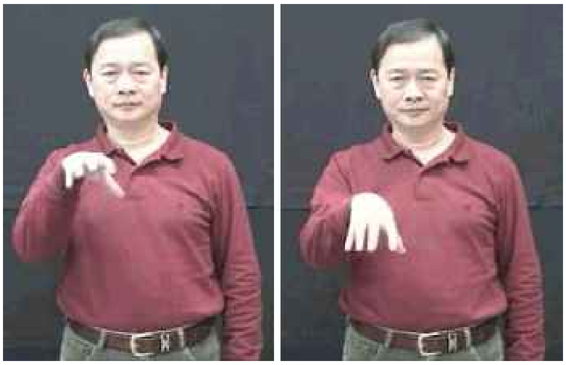
ZERO ([*ZERO*])

HAVE ([*HAND*])

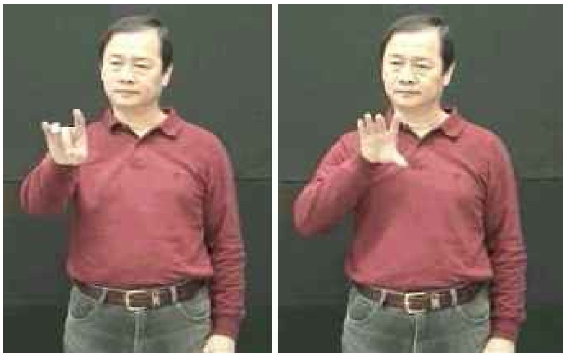
SIX ([*SIX*])

LÜ ([*LÜ*] on both hands)

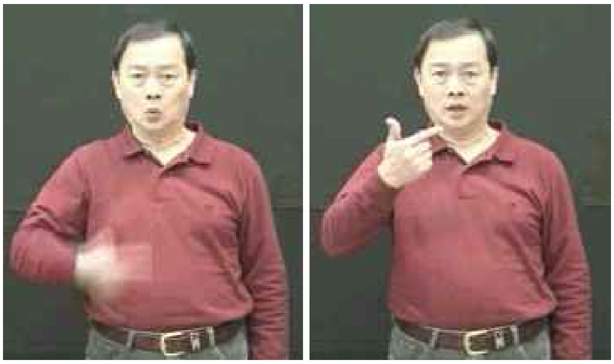
PLACE ([*SAME*])



RENT (handshape changes from [*RENT*] to [*SAME*])

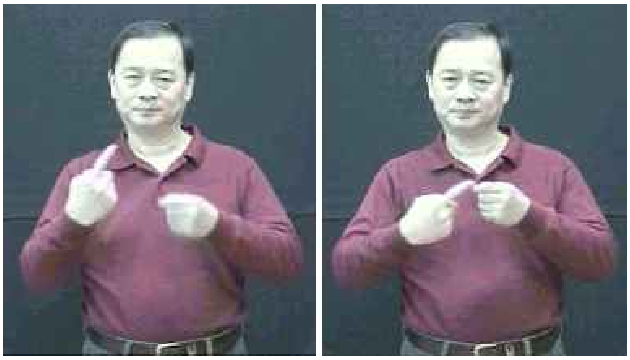


FAST ([*SIX*])



CUT CLASS ([*ZERO*] on dominant hand, [*HAND*] on non-dominant hand)

RICE ([*ONE*] on dominant hand, [*LÜ*] on non-dominant hand)



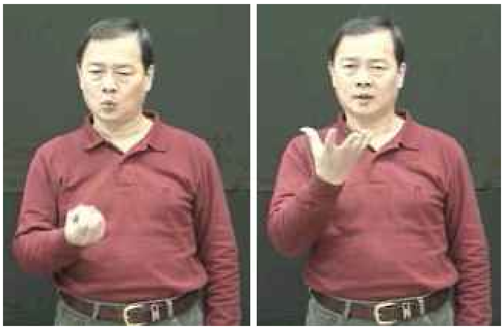
WRITE ([*LÜ*] on dominant hand, [*HAND*] on non-dominant hand; dominant hand moves downward across the non-dominant hand as if writing)

STICK (dominant hand changes from [*RENT*] to [*HAND*]; non-dominant hand remains [*HAND*] throughout)

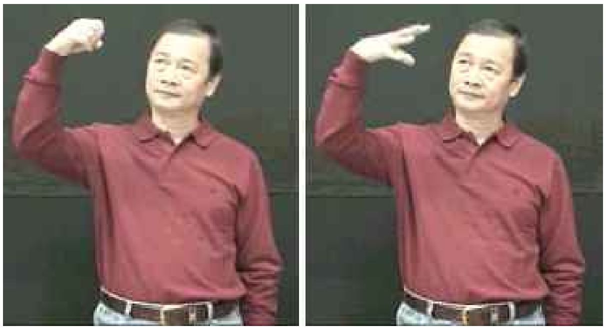
Appendix B: The nine signs involving handshape change used as production targets in the experiment, grouped by the sets that formed the basis of the experimental design.

Set 1: [*ZERO*] > [*SAME*]

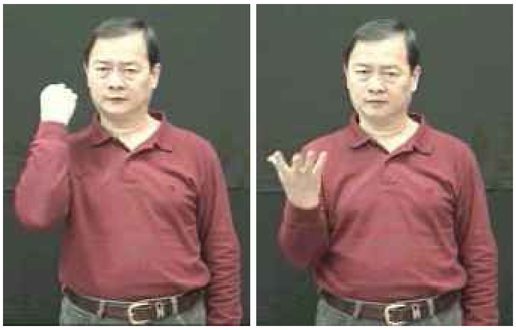
FLOWER



SUN



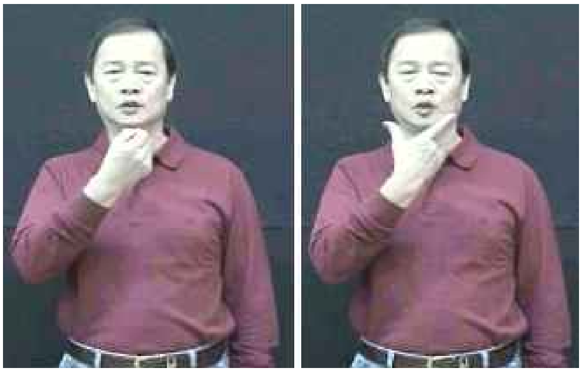
NEW



Set 2: [*LÜ*] > [*SIX*]

SMART

BEAN



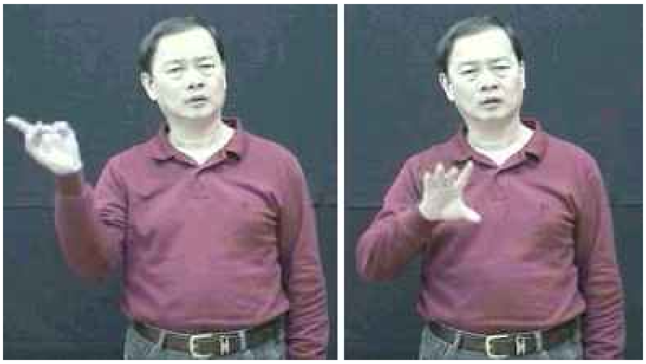
WAKE UP

Set 3: [*RENT*] > [*SAME*]

NO BIG DEAL

INVENT

NEVER BEFORE



References

Abdel Rahman, Rasha, Miranda van Turennout, and Willem J. M. Levelt. 2003. Phonological encoding is not contingent on semantic feature retrieval: an electrophysiological study on object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29.5:850-860.

Armstrong, D., W. Stokoe, and S. Wilcox. 1995. *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.

Blair, I. V., G. R. Urland, and J. E. Ma. 2002. Using internet search engines to estimate word frequency. *Behavior Research Methods, Instruments, and Computers* 34.2: 286-290.

Boersma, Paul. 1998. *Functional Phonology.* The Hague: Holland Academic Graphics.

Brentari, Diane. 1990. Licensing in ASL handshape change. *Sign Language Research: Theoretical Issues*, ed. by Ceil Lucas, 27-49. Washington, DC: Gallaudet University Press.

Brentari, Diane. 1998. *A Prosodic Model of Sign Language Phonology*. Cambridge: MIT Press.

Channon, Rachel. 2002. *Signs Are Single Segments: Phonological Representations and Temporal Sequencing in ASL and Other Sign Languages.* College Park: University of Maryland dissertation.

Chen, J.-Y., T.-M. Chen, and G. Dell. 2002. Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language* 46: 751-781.

Chen, J.-Y., W.-C. Lin, and L. Ferrand. 2003. Masked priming of the syllable in Mandarin Chinese speech production. *Chinese Journal of Psychology* 45:107-120.

Cholin, Joana, Niels O. Schiller, and Willem J. M. Levelt. 2004. The preparation of syllables in speech production. *Journal of Memory and Language* 50:47-61.

Corina, David P. 1990. Handshape assimilations in hierarchical phonological representation. *Sign Language Research: Theoretical Issues*, ed. by Ceil Lucas, 27-49. Washington, DC: Gallaudet University Press.

Corina, David P. 1993. To branch or not to branch: underspecification in ASL handshape contours. *Current Issues in ASL Phonology*, Vol. 3: *Phonetics and Phonology*, ed. by Geoffrey Coulter, 63-95. New York: Academic Press.

Corina, David P., and Ursula C. Hildebrandt. 2002. Psycholinguistic investigations of phonological structure in ASL. *Modality and Structure in Signed and Spoken Languages*, ed. by R. P. Meier, K. Cormier, and D. Quinto-Pozos, 88-111. Cambridge: Cambridge University Press.

Coulter, Geoffrey R. (ed.) 1993. *Current Issues in ASL Phonology*. New York: Academic Press.

Cutler, Anne. (ed.) 1982. *Slips of the Tongue and Language Production*. The Hague: Mouton.

Fromkin, Victoria A. 1971. The non-anomalous nature of anomalous utterances. *Language* 47:27-52.

Fromkin, Victoria A. 1973. *Speech Errors as Linguistic Evidence*. The Hague: Mouton.

Fromkin, Victoria A. (ed.) 1980. *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand*. New York: Academic Press.

Garrett, M. F. 1980. The limits of accommodation. *Errors in Linguistic Performance*, ed. by V. A. Fromkin, 263-271. New York: Academic Press.

Garrett, M. F. 1988. Processes in language production. *Language: Psychological and Biological Aspects*, ed. by F. J. Newmeyer, 69-96. Linguistics: The Cambridge Survey, Vol. 3. Cambridge: Cambridge University Press.

Hildebrandt, Ursula C., and David P. Corina. 2002. Phonological similarity in American Sign Language. *Language and Cognitive Processes* 17:593-612.

Hohenberger, Annette, Daniela Happ, and Helen Leuninger. 2002. Modality-dependent aspects of sign language production: evidence from slips of the hands and their repairs in German Sign Language. *Modality and Structure in Signed and Spoken Languages*, ed. by R. P. Meier, K. Cormier, and D. Quinto-Pozos, 112-142. Cambridge: Cambridge University Press.

Kirchner, Robert. 1997. Contrastiveness and faithfulness. *Phonology* 14:83-111.

Klima, Edward S., and Ursula Bellugi. 1979. *The Signs of Language*. Cambridge: Harvard University Press.

Levelt, Willem J. M., and Peter Indefrey. 2000. The speaking mind/brain: where do spoken words come from? *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium*, ed. by Alec Marantz, Yasushi Miyashita, and Wayne O'Neil, 77-93. Cambridge: MIT Press.

Levelt, Willem J. M., Peter Praamstra, Antje S. Meyer, Paivi Helenius, and Riitta Salmelin. 1998. An MEG study of picture naming. *Journal of Cognitive Neuroscience* 10:553-567.

Levelt, Willem J. M., Ardi Roelofs, and Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22:1-75.

Levelt, Willem J. M., and Linda Wheeldon. 1994. Do speakers have access to a mental syllabary? *Cognition* 50:239-269.

Liberman, Alvin M. 1996. *Speech: A Special Code*. Cambridge: MIT Press.

Liddell, Scott K. 1990. Structures for representing handshape and local movement at the phonemic level. *Theoretical Issues in Sign Language Research*, Vol. 1: *Linguistics*, ed. by Susan Fischer and Patricia Siple, 37-65. Chicago: University of Chicago

Press.
Liddell, Scott K., and Robert E. Johnson. 1989. American Sign Language: the phonological base. *Sign Language Studies* 64:195-277.
Lombardi, L. 1990. The nonlinear representation of the affricate. *Natural Language and Linguistic Theory* 8:375-425.
Lucas, Ceil, and Clayton Valli. (eds.) 2000. *Linguistics of American Sign Language* (3rd edition). Washington, DC: Gallaudet University Press.
Mayberry, R. I., and S. D. Fischer. 1989. Looking through phonological shape to lexical meaning: the bottleneck of non-native sign language processing. *Memory and Cognition* 17:740-754.
Meyer, Antje S. 1990. The time course of phonological encoding in language production: the encoding of successive syllables of a word. *Journal of Memory and Language* 29:524-545.
Meyer, Antje S. 1991. The time course of phonological encoding in language production: phonological encoding inside a syllable. *Journal of Memory and Language* 30:69-89.
Meyer, Antje S., and H. Schriefers. 1991. Phonological facilitation in picture-word interference experiments: effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Language, Memory, and Cognition* 17:1146-1160.
Moy, Anthony. 1990. A psycholinguistic approach to categorizing handshapes in American Sign Language: is [A_S] an allophone of /A/? *Sign Language Research: Theoretical Issues*, ed. by Ceil Lucas, 346-357. Washington, DC: Gallaudet University Press.
Myers, James, and Jane Tsay. 2003. A formal functional model of tone. *Language and Linguistics* 4.1:105-138.
Newkirk, Don, Edward S. Klima, Carlene C. Pedersen, and Ursula Bellugi. 1980. Linguistic evidence from slips of the hand. *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand*, ed. by Victoria A. Fromkin, 165-197. New York: Academic Press.
Padden, Carol A., and David M. Perlmutter. 1987. American Sign Language and the architecture of phonological theory. *Natural Language and Linguistic Theory* 5.3:335-375.
Roelofs, Ardi. 1997. The WEAVER model of word-form encoding in speech production. *Cognition* 64:249-284.
Roelofs, Ardi. 1999. Phonological segments and features as planning units in speech production. *Language and Cognitive Processes* 14:173-200.
Roelofs, A., and A. S. Meyer. 1998. Metrical structure in planning the production of

spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24:922-939.

Sandler, Wendy. 1989. *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language*. Dordrecht: Foris.

Sandler, Wendy. 1990. Temporal aspects and ASL phonology. *Theoretical Issues in Sign Language Research*, ed. by Susan D. Fischer and Patricia Siple, 7-35. Chicago: University of Chicago Press.

Sandler, Wendy. 2000. One phonology or two? Sign language and phonological theory. *The First Glot International State-of-the-Article Book: The Latest in Linguistics*, ed. by Lisa Chen and Rint Sybesma, 349-383. Berlin: Mouton de Gruyter.

Schneider, W., A. Eschman, and A. Zuccolotto. 2002. *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools Inc.

Schriefers, H., A. S. Meyer, and W. J. M. Levelt. 1990. Exploring the time course of lexical access in language production: picture-word interference studies. *Journal of Memory and Language* 29:86-102.

Smith, Wayne. 1989. *The Morphological Characteristics of Verbs in Taiwan Sign Language.* Bloomington: Indiana University dissertation.

Smith, Wayne H., and Li-fen Ting. 1979. *Shou Neng Sheng Qiao* [*Your Hands Can Become a Bridge*], Vol. 1. Taipei: Deaf Sign Language Research Association of the Republic of China.

Smith, Wayne H., and Li-fen Ting. 1984. *Shou Neng Sheng Qiao* [*Your Hands Can Become a Bridge*], Vol. 2. Taipei: Deaf Sign Language Research Association of the Republic of China.

Stemberger, J. P. 1983. *Speech Errors and Theoretical Phonology: A Review.* Bloomington: Indiana University Linguistics Club.

Steriade, Donca. 1993. Closure, release, and nasal contours. *Nasals, Nasalization and the Velum*, ed. by M. Huffman and R. Krakow, 125-153. Phonetics and Phonology 5. New York: Academic Press.

Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. *Papers in Laboratory Phonology*, Vol. 5: *Acquisition and the Lexicon*, ed. by M. B. Broe and J. B. Pierrehumbert, 313-334. Cambridge: Cambridge University Press.

Stokoe, William C. 1960. *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf*. Studies in Linguistics, Occasional Papers 8. Buffalo: Dept. of Anthropology and Linguistics, University of Buffalo.

Stokoe, William C., Dorothy C. Casterline, and Carl G. Croneberg. 1965. *A Dictionary of American Sign Language on Linguistic Principles*. Silver Spring: Linstok Press.

Taub, S. 2000. *Language and the Body: Iconicity and Metaphor in American Sign Language*. Cambridge: Cambridge University Press.

Uyechi, Linda. 1996. *The Geometry of Visual Phonology*. Stanford: CSLI.

van der Hulst, Harry, and Anne Mills. (eds.) 1996. *Sign Linguistics: Phonetics, Phonology and Morpho-syntax. Lingua* 98 (special issue).

Wheeldon, L., and W. J. M. Levelt. 1995. Monitoring the time course of phonological encoding. *Journal of Memory and Language* 34:311-334.

Whittemore, Gregory L. 1987. *The Production of ASL Signs*. Austin: University of Texas dissertation.

Winer, B. J. 1971. *Statistical Principles in Experimental Design*. New York: McGraw-Hill Book Company.

[Received 15 September 2003; revised 18 April 2004; accepted 20 July 2004]

James Myers
Graduate Institute of Linguistics
National Chung Cheng University
Minhsiung, Chiayi 621, Taiwan
Lngmyers@ccu.edu.tw

台灣手語的音韻產生歷程

麥　傑　　李信賢　　蔡素娟
國立中正大學

這篇論文主要在描述一個有關台灣手語在語言產生過程中的手型變化（相當於口語的音韻變化）的實驗。這個實驗採用的是「隱藏啓動實驗典範」(implicit priming; Meyer 1990, 1991)。實驗結果不僅提供證據，證明音韻形式在手語的產生中扮演一個重要的角色，而且顯示手語音韻產生的時程與 Levelt, Roelofs, and Meyer 等人 (1999) 所發現的口語產生的時程一致。本實驗更進一步凸顯有關音韻產生（包括手語與口語）的研究方法的某些考量的重要性。

關鍵詞：台灣手語，音韻學，心理語言學